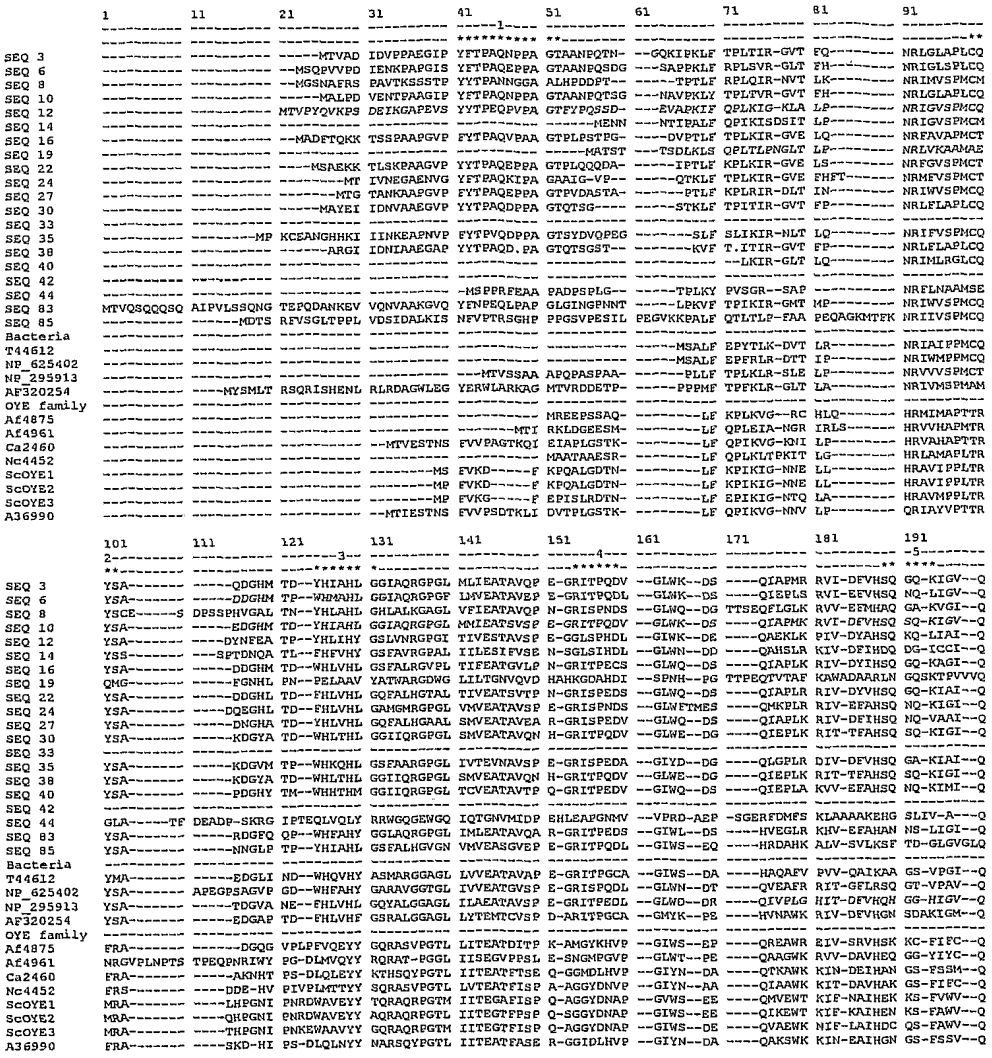

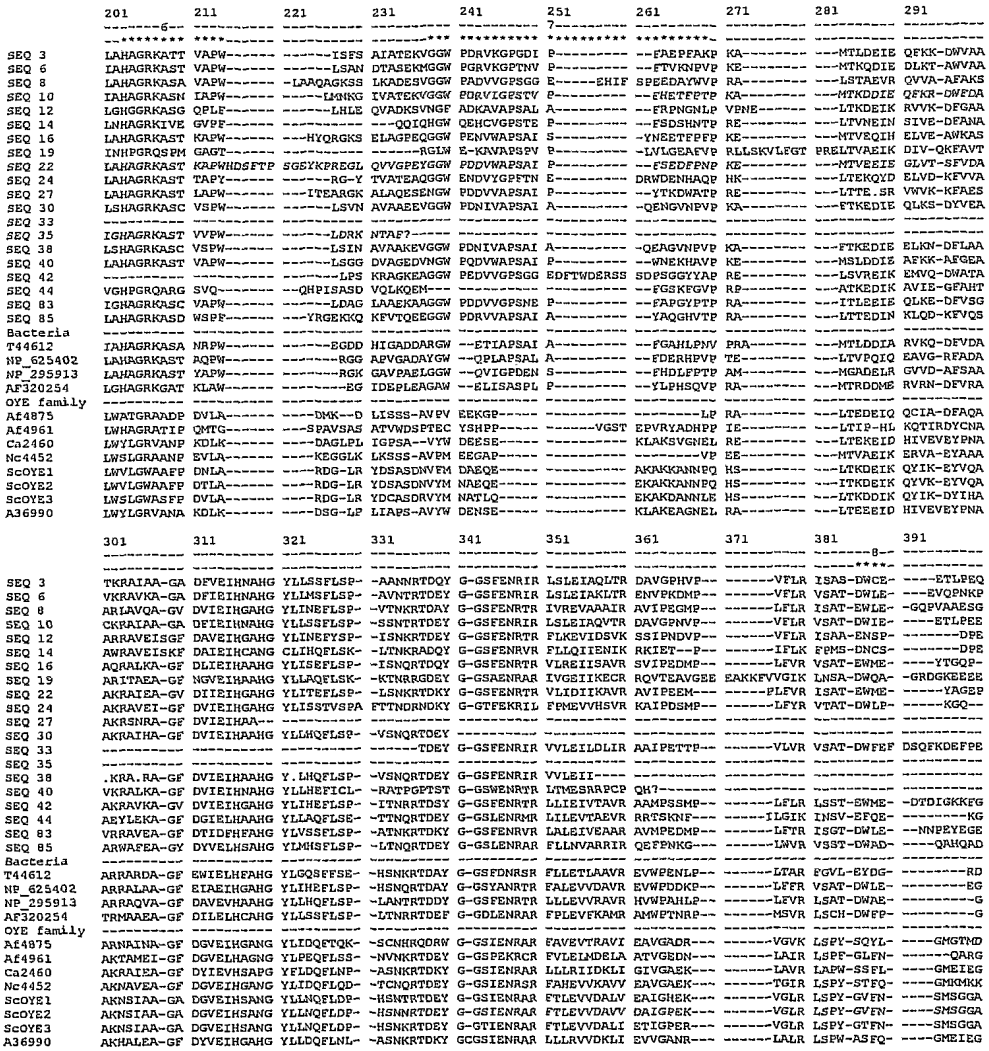

```
                401         411         421         431         441         451         461         471         481         491
                ----------  ----------  -------9--- ----------  10--------  ----------  ----------  ----------  ----------  ----------
                ----------  ----------  ---******* ##########  ##########  ##----------  ----------  ----------  ----------  ----------
SEQ 3           SWKSEDTVR-  FAQELVK--Q  GAVDLIDISS  GGVLAQQ---  ---------KI  KSGPAFQVPF  AVAVKKAVGD  ---KLLVAAV  GAIT------  ----------
SEQ 6           SWRGVDTVR-  -FAKILA-ET  GYVDVLDVSS  GGTHSEQ---  ---------HI  HAKPGFQAPF  AIAVKNAVGD  ---KLAVASV  GMIA------  ----------
SEQ 8           SWDM-QSSL-  ELVKKLP--E  WGIDLVDVSS  AANHKDQ---  ---------KI  NLHTAYQTDL  AGQIRQAI--  ---RAAGAST  LVGAVGLITD  SEQARGLVQG
SEQ 10          SWKLSDSVR-  FAEALAA--Q  GAIDLIDVSS  GGVHAAQ---  ---------KI  KSGPAFQAPF  AVAIKKAVGD  ---KLLVATV  GTIT------  ----------
SEQ 12          AWTIEDSKK-  -LADILV--E  KGIALVDVSS  GGNDYRQPP-  ----RSGISK  ELREPIHVPL  SRAIKQHVGD  ---KLLVSCV  GGLE------  ----------
SEQ 14          AWSTEDALK-  -LADLVI--D  LGVKVIDVTS  GGNVAHCKS-  ----RYLLND  DKQLPSQVPL  ARKLKSHIRN  ---RCLIACS  GGLD------  ----------
SEQ 16          SWDLQQTI--  ELAKILP--D  LGVDLLDVSS  GGNNKDQ---  ---------KI  NVHTYYQIDM  AEQIRAAVHE  AGKQLLVGAV  GLVT----SA  EIAKETVQEK
SEQ 19          TDTAEEVLK-  -QIELFE--Q  WGIDFVEVSG  GSYEDPQMAN  GPKPEKSERT  MAREAFLEF  AKIIRTK---  -FPKLPLMVT  GGFR------  ----------
SEQ 22          SWDLEQSTQ-  -LAKLLP--D  LGVDLLDVSS  GGNSVAQ---  ---------KI  ELTPYYQIDL  AAKIREAVGD  ---RLLIGAV  GNIN------  ----------
SEQ 24          GWEIEDTVAF  TLAARLR--D  GGVDLIDVSS  GGNHKDQ---  ---------RI  EVKDCYQVPF  AEKIKDQVNG  ----ILLGAV  GMIR------  ----------
SEQ 27          ----------  ----------  ----------  ----------  ----------  ----------  ----------  ----------  ----------  ----------
SEQ 30          ----------  ----------  ----------  ----------  ----------  ----------  ----------  ----------  ----------  ----------
SEQ 33          SWTVEQTC--  QLARILP--K  HGVDLVDVSS  GGIHPKS---  ------AIAI  KSGPAYQVDL  AKQVKKAVGD  ---SVLVSAV  GGIK------  ----------
SEQ 35          ----------  ----------  ----------  ----------  ----------  ----------  ----------  ----------  ----------  ----------
SEQ 38          ----------  ----------  ----------  ----------  ----------  ----------  ----------  ----------  ----------  ----------
SEQ 40          ----------  ----------  ----------  ----------  ----------  ----------  ----------  ----------  ----------  ----------
SEQ 42          SWDVESTIK-  -ISKILA--D  LGVDLLDVSS  GGNHPQQ---  ---------KI  NMENT-----  ----------  ----------  ----------  ----------
SEQ 44          FKP-EEAVQ-  LCEALEAAGM  ---DPVETSG  GTYESFG---  --FAHRKESS  RKRENYFIEF  AEVIRKAVKH  ---MVVYTTG  GFKT------  ----------
SEQ 83          TWTLEQSIK-  -LAHQLA--D  RGVDVLDVSS  GGIHKMQ---  ---------KV  AAGPGYQAPL  AKAIKKSVGD  ---KMLISTV  GSIK------  ----------
SEQ 85          SWTVDQTVE-  LAKMLQE---  ARVDLLDVSS  GGLVPFQ---  ---------KI  TVGAGYQLEG  AKAVRDALAK  --IEPDASKR  MLVGA-----  ----------
Bacteria        ----------  ----------  ----------  ----------  ----------  ----------  ----------  ----------  ----------  ----------
T44612          EQTLEESI--  ELARRFK--A  GGLDLLSVSV  GFTIPET---  ---------NI  PWGPAFMGPI  AERVRREAKL  -----PVTSAW  GFGT------  ----------
NP_625402       GWTPDDTVR-  -FARDLE--A  HGIDLLDVST  GGNVPRV---  ---------RI  PTGPGYQVPF  AARVKAGST-  ----LPVAAV  GLIT------  ----------
NP_295913       GWDLEQTVQ-  -LSKLLK--Y  EGVDVLDISS  GGLTAAQ---  ---------QI  EVGPGYQVPF  AAAVSRAETE  ----ISVMAV  GLIE------  ----------
AF320254        GNTADDAVA-  -IARLFK--E  AGADIIDCSS  GQVWKGD---  ---------QP  VYGRMYQTPF  ADRIRNEVGI  ----PTLAVG  AISE------  ----------
OYE family      ----------  ----------  ----------  ----------  ----------  ----------  ----------  ----------  ----------  ----------
Af4875          EL--VPQFEY  LIA-----QM  RRLDVAYLHL  ANSRWL----  ----------DE  EKPHPDPNHE  VFVRVWG-Q-  --SS-PILLA  GGYD------  ----------
Af4961          EQR-VETWTF  LCESLKKAHP  ---NLSYVSF  IEPRYE----  ----------  -QIFSYEEKD  NFLRSWG---  LSDVDLSSFR  KIFGTTPFFS  ----------
Ca2460          EE-----IHSY  ILQQLQQRAD  NGQQLAYVSL  IEPRVIG---  ----IFDASL  EDQKGRSNEF  AYKYWKG---  -----NFVRA  GNYT------  ----------
Nc4452          DLIP--QFED  VIRKIN----  -GFGLAYLHL  TQSRVAGN---  -----MDVQP  EEDEE-NLAF  AAKLWDG---  -----PLLIA  GGLT------  ----------
ScOYE1          ETGIVAQYAY  VAGELEKRAK  AGKRLAFVHL  VEPRVTNP--  ----FLTEGE  GEYEGGSNDF  VYSIWKG---  -----PVIRA  GNFA------  ----------
ScOYE2          ETGIVAQYAY  VLGELERRAK  AGKRLAFVHL  VEPRVTNP--  ----FLTEGE  GEYNGGSNKF  AYSIWKG---  -----PIIRA  GNFA------  ----------
ScOYE3          EPGIIAQYSY  VLGELEKRAK  AGKRLAFVHL  VEPRVTDP--  ----SLVEGE  GEYSEGTNDF  AYSIWKG---  -----PIIRA  GNYA------  ----------
A36990          EE----IHSY  ILQQLQQRAD  NGQQLAYISL  VEPRVTG---  ----IYDVSL  KDQQGRSNEF  AYKIWKG---  -----NFIRA  GNYT------  ----------

                501         511         521         531         541         551         561         571         581         591
                ----------  ----------  ----------  ----------  ----------  ----------  ----------  ----------  ----------i 1---------
                ----------  ----------  ----------  ----------  ----------  ----------  ----------  ----------  ----------**** ***-------
SEQ 3           --NGKQ--AN  QILEEQD---  ----------  ----------  ---IDVALVG  RGFQKDPGLA  WTFAQHLGV-  ----------  -----EISMAN  QIRWGFTRRG
SEQ 6           ---SAHLANS  LLEKDG----  ----------  ----------  ---LDLVLVG  RGFQKNPGLV  WAWADELNV-  ----------  -----EISMAN  QIRWGFSRRG
SEQ 8           ADEATAAEAM  LSGPEPK---  ----------  ----------  ---ADAILIA  RQFLREPEWV  FSTARKLGV-  ----------  -----PVTVPV  QFGRAI----
SEQ 10          --NGKQ--AN  KLLEEEG---  ----------  ----------  ---LDVALVG  RGFQKDPGLA  WTFAQHLDV-  ----------  -----EIAMAS  QIRWGFTRRG
SEQ 12          ---KDPELLN  KYLEEGT---  ----------  ----------  ---FDLALIG  RGFLRNPGLV  WEFADKLGV-  ----------  -----RLHQAL  QLGWGFWPNK
SEQ 14          ---RDIFKLD  EFIANGD---  ----------  ----------  ---FDIALIG  KGFLKNTGLI  SRIADQLQA-  ----------  -----QFRTAP  QYKLALS---
SEQ 16          E-DGRVTIQR  ENGAKTR---  ----------  ----------  ---ADMVLVA  RQFLKEPEFV  LTVADELGV-  ----------  -----DVKAPV  QYLRGPLSSR
SEQ 19          ----TRQGME  AALESDD---  ----------  ----------  ---CDMIGIG  RPAIINPSLP  ANLILNPEV-  ----------  -----PDADAR  LFDKKRAEPH
SEQ 22          --TADI--AR  DVVDEQGAEK  VAEAKQTHDT  IEVVSESHGG  KTKADLVLIA  RQFLREPEFV  LRTAHNLGV-  ----------  -----NVQWPH  QYHRAVWRKG
SEQ 24          --DGLFTTAN  EILESGK---  ----------  ----------  ---ADVTFVA  REFLRNPSLV  LDSANQLGE-  ----------  -----NVAWPV  QYDYAVKGHR
SEQ 27          ----------  ----------  ----------  ----------  ----------  ----------  ----------  ----------  ----------  ----------
SEQ 30          ----------  ----------  ----------  ----------  ----------  ----------  ----------  ----------  ----------  ----------
SEQ 33          --TGHL--AE  EVLQSG----  ----------  ----------  ---IDIVRAG  RWEQQNPGLV  RAFANELGV-  ----------  -----EVKMAN  QIDWSFKGRG
SEQ 35          ----------  ----------  ----------  ----------  ----------  ----------  ----------  ----------  ----------  ----------
SEQ 38          ----------  ----------  ----------  ----------  ----------  ----------  ----------  ----------  ----------  ----------
SEQ 40          ----------  ----------  ----------  ----------  ----------  ----------  ----------  ----------  ----------  ----------
SEQ 42          ----------  ----------  ----------  ----------  ----------  ----------  ----------  ----------  ----------  ----------
SEQ 44          --VGAM-VDA  LQGVDG----  ----------  ----------  ------IGIG  RAAGSEPDLA  KDIIAGKVSS  IIKYAMGEDE  FVLQLTACSA  QIRLMAKGEE
SEQ 83          --IGTL--AE  EIIAGG----  ----------  --------ED  DTPLDLVASG  RLFQKNTGLV  WSWADDLNT-  ----------  -----SIQIAH  QIAWGFGGRA
SEQ 85          --VGMM--EG  SYDSPNG---  ----------  ----------  ---QDRSQIG  KLAEQSIQSG  ECDAVLLAR-  ---GLMSYPS  WTEDASVALM  GTRAAGNPQY
Bacteria        ----------  ----------  ----------  ----------  ----------  ----------  ----------  ----------  ----------  ----------
T44612          -PQLAE----  AALQANQ---  ----------  ----------  ---LDLVSVG  RAHLADPHWA  YFAAKELGV-  ----------  -----EKASWT  LPAPYAHWLE
NP_625402       -EPG---QAE  KILANGE---  ----------  ----------  ---ADAVLLG  RELLRNPSWA  QHAARELGV-  ----------  -----DARMPD  QYGWGM----
NP_295913       --TGA--QAE  AILQAGD---  ----------  ----------  ---ADLIALG  RPFLRDPHWA  QRAARELGL-  ----------  -----RPVSID  QYARAGW---
AF320254        ---AD--HAN  SIIAAGR---  ----------  ----------  ---ADLCAIA  RPHLADPAWT  LHEAAKIGF-  ----------  -----GEVAWP  KQYRSARGQY
OYE family      ----------  ----------  ----------  ----------  ----------  ----------  ----------  ----------  ----------  ----------
Af4875          AASAEKVTEQ  MAAATYT---  ----------  ----------  ---NVAIAFG  RYFISTPDLP  FRVMAGIQL-  ----------  -----QKYDRA  SFYSTLSREG
Af4961          AGGWDQSNSW  GVLEEGR---  ----------  ----------  ---YDALLYG  RYFTSNPDLV  ERLRKGIPF-  ----------  -----TPYDRS  RFYGPFEDNA
Ca2460          -YDAPEFKTL  LHDLDND---  ----------  ----------  ---RTIVGEA  RFFTSNPDLV  EKLKLGKPL-  ----------  -----NHYDRE  EFYKYYNY-G
Nc4452          -PETAK-HLV  DREFPEK---  ----------  ----------  ---DVVATFG  RHFISTPDLP  FRIKEGIEL-  ----------  -----NPYDRD  TFYKAKSPDG
ScOYE1          -LHP---EVV  REEVKDK---  ----------  ----------  ---RTLIGYG  RFFISNPDLV  DRLEKGLPL-  ----------  -----NKYDRD  TFYQMSAH-G
ScOYE2          -LHP---EVV  REEVKDP---  ----------  ----------  ---RTLIGYG  RFFISNPDLV  DRLEKGLPL-  ----------  -----NKYDRD  TFYKMSAE-G
ScOYE3          -LHP---EVV  REQVKDP---  ----------  ----------  ---RTLIGYG  RFFISNPDLV  YRLEEGLPL-  ----------  -----NKYDRS  TFYTMSAE-G
A36990          -YDAPEFKTL  INDLKND---  ----------  ----------  ---RSIIGFS  RFFTSNPDLV  EKLKLGKPL-  ----------  -----NYYNAE  EFYKYYNY-G
```

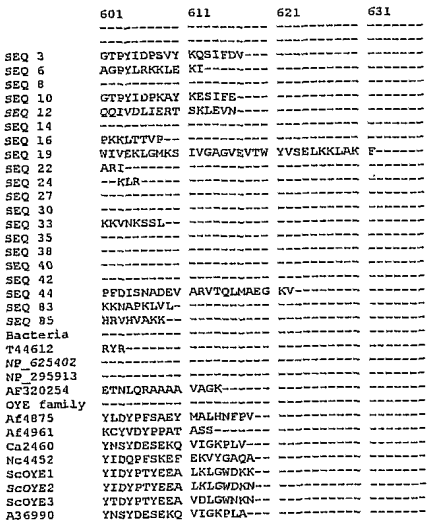

Figure 1. A multiple alignment of the 2031 OR amino acid sequence from *A. fumigatus* (SEQ ID No3) along with related 2031 ORs from other fungi and bacteria (see Example 4) and OYEs. Regions 1-11, marked with * or #, refer to amino acids conserved between ORs but not OYEs.

Fungal 2031 ORs are given by the following SEQ ID No.: *A. fumigatus*, SEQ ID Nos. 3, 6 and 8; *A. nidulans*, SEQ ID No. 10; *C. albicans* SEQ ID Nos. 12 and 14; *N. crassa*, SEQ ID Nos. 16 and 19; *M. grisea* SEQ ID Nos. 22 and 44; *S. pombe* SEQ ID No. 24 (NP_595868); *C. trifolii* SEQ ID No. 27; *F. sporotrichioides* SEQ ID Nos. 30, 33 and 35; *F. graminearum* SEQ ID Nos. 38 and 83; *M. graminicola* SEQ ID Nos. 40 and 42; *U. maydis* SEQ ID No 85.

Bacterial ORs resembling 2031 are:
T44612 (*Pseudomonas putida*), SEQ ID No. 86; NP_625402 (*Streptomyces coelicolor*), SEQ ID No. 87; NP_295913 (*Deinococcus radiodurans*), SEQ ID No. 88; AF320254 (*Azoarcus evansii*, SEQ ID No. 89.

Fungal ORs similar to the Old Yellow Enzyme family (originally identified in *S. cerevisiae*):
A. fumigatus, Af4875 and Af4961, SEQ ID Nos. 90 and 91 respectively; *C. albicans*, Ca2460 and A36990, SEQ ID Nos. 92 and 93 respectively; *N. crassa*, Nc4452, SEQ ID No. 94; *S. cerevisiae*, OYE1, OYE2 and OYE3, SEQ ID Nos. 95-97 respectively.

Details of the sequence searches that identified the ORs other than SEQ ID No. 3, and methods for the construction of multiple alignments are given in Example 4 hereinafter.

```
          1          11         21         31         41         51         61         71         81         91
          ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
          ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 1     GTTCGACGTC ATTGCCACGT TTCGACCCAA GGGCAGACGC CATGTCGCCG AGCGATCGCC GCGATATGCC TCGAATTTGC GCCATTCGGC ATCCAGTTTC
SEQ 2     ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 4     ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 5     ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 7     ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 9     ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 11    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 13    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 15    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 17    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 18    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 20    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 21    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 23    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 25    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 26    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 28    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 29    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 32    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 34    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 36    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 37    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 39    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 41    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 43    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 82    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 84    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------

          101        111        121        131        141        151        161        171        181        191
          ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
          ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 1     CAGTGCCCTT CCCCGAATGA CTGTCTCCAC TATTCGGCAA GATTGTAAAT CAAGCCTGAA GAAGCGGAGC AATTCTTGGA AGTCGTATGT TCTACTGATT
SEQ 2     ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----GTATGT TCTACTGATT
SEQ 4     ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 5     ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 7     ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 9     ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 11    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 13    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 15    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 17    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 18    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 20    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 21    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 23    ---------- ---------- ---------- ---------- ---------- ---------- ---------- -----CGAAA CCTCGACCCA AACAAACAGC
SEQ 25    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 26    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ------GAAC
SEQ 28    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 29    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 32    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 34    ----AGGAAG TTGCATGTCA CTTGTAGTGA CAGGGCGTCG TGTAAATTTT ATAAATACCT ATACTTGTTT GTTCACTTCT ATGCTACTCA TATCAATCCG
SEQ 36    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 37    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 39    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 41    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 43    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 82    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 84    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------

          201        211        221        231        241        251        261        271        281        291
          ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
          ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 1     TCTGTGCCTG GCGCAGACGG GTATATAAAT AAAGATCACC GCACCGAGGA GTTTCTTACC AACCCATCAA TAACCATCCA CAATCTCCTA CAACAAAAAT
SEQ 2     TCTGTGCCTG GCGCAGACGG GTATATAAAT AAAGATCACC GCACCGAGGA GTTTCTTACC AACCCATCAA TAACCATCCA CAATCTCCTA CAACAAAAAT
SEQ 4     ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------A TGTCGCAACC
SEQ 5     ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------A TGTCGCAACC
SEQ 7     ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------A TGGGTTCCAA
SEQ 9     ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- --------AT
SEQ 11    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---ATGACAG TTCCATACCA
SEQ 13    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------A TGGCCGACTT
SEQ 15    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 17    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 18    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- -----ATGTC
SEQ 20    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- -----ATGTC
SEQ 21    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 23    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 25    TGACCCTCTC CTTGACAACA AAGCCGGCCA TCCTCGCCGA CGATTGCCTC TACCCCCGCA TAGTCACACT CGCACGTCCG TTCTCCCACC GTCAAACAGA
SEQ 26    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 28    TGCTGTAGAT GTGGTTGAAT TGGTATATTA GACCGGAGTA CTCTATATGC GAGAGACTAT ACATTGAAGT TGCCAACGTT CTTCCAGATT GATTAATCAT
SEQ 29    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- --------AT
SEQ 32    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 34    AGAAGATCAA ACAGTCCCCT ATACACACTT GTCAAGACCT ATCTATTATT TCAAAAATCA GCAATATGGC TGAGACAATG CCTAAGTGTG AGGCAAATGG
SEQ 36    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 37    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 39    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 41    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 43    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 82    ---------- ----------TGACAG TTCAATCACA GCAACAATCC CAGGCTATTC CCGTCCTTTC TTCCCAGAAT GGCACTGAAC CCCAAGACGC
SEQ 84    ---------- ---------- ---------- ---------- ---------- ---------- --------AT GGACACGTCT CGATTCGTGT CTGGTCTCAC
```

```
          301        311        321        331        341        351        361        371        381        391
          ---------- ---------- ---------- ---------- ---------- ---------- -1-------- ---------- ---------- ----------
          ---------- ---------- ---------- ---------- ---******* ********** ********** ********** **-------- ----------
SEQ 1     GACTGTCGCC GATATCGACG TTCCTCCTGC CGAGGGCATC CCCTACTTCA CTCCGGCCCA GAACCCTCCT GCCGGTACGG CAGCTAACCC CCAGACCAAT
SEQ 2     GACTGTCGCC GATATCGACG TTCCTCCTGC CGAGGGCATC CCCTACTTCA CTCCGGCCCA GAACCCTCCT GCCGGTACGG CAGCTAACCC CCAGACCAAT
SEQ 4     TGTTGTGCCT GACATCGAGA ACAAACCCGC GCCGGGTATC TCGTACTTTA CTCCGGCGCA AGAGCCGCCT GCTGGCACCG CTGCTAATCC TCAGTCTGAT
SEQ 5     TGTTGTGCCT GACATCGAGA ACAAACCCGC GCCGGGTATC TCGTACTTTA CTCCGGCGCA AGAGCCGCCT GCTGGCACCG CTGCTAATCC TCAGTCTGAT
SEQ 7     CGCCTTCCGG TCCCCCGCCG TCACCAAGTC CTCCTCCACC CCCTACTACA CTCCCGCCAA CAATGGAGGC GCCGCCCTGC ACCCCGACGA CCCCAC----
SEQ 9     GGCTCTCCCT GACGTCGAAA ACACCCCCGC CGCCGGCATC CCCTACTTTA CACCAGCACA GAACCCTCCT GCTGGAACAG CTGCCAACCC GCAAACCAGC
SEQ 11    AGTAAAACCA TCAGATGAAA TCAAAGGTGC TCCTGAGGTT TCCTATTACA CTCCAGAACA GCCTGTTCCG GCTGGTACTT TTTATCCCCA ATCGTC----A
SEQ 13    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---ATGGAAA ACAACAATAC TATACCG---
SEQ 15    CACCCAGAAG AAGACCTCCT CCCCCGCGGC CCCGGGTGTT CCCTTCTACA CCCCGGCCCA GGTCCCCGCC GCCGGCACTC CCCTCCCCTC CACCCCC---
SEQ 17    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ATGGCTACTT CCACTACCTC CGACCTC---
SEQ 18    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ATGGCTACTT CCACTACCTC CGACCTC---
SEQ 20    GGCAGAAAAG AAGACTTTGA GCAAACCGGC CGCCGGGGTG CCTTACTACA CCCCAGCCCA GGAGCCGCCG GCAGGGACCC CTTTGCAGCA GCAGGACG--
SEQ 21    GGCAGAAAAG AAGACTTTGA GCAAACCGGC CGCCGGGGTG CCTTACTACA CCCCAGCCCA GGAGCCGCCG GCAGGGACCC CTTTGCAGCA GCAGGACG--
SEQ 23    -----ATGAC TATTGTTAAT GAAGGAGCCG AAAATGTTGG TTATTTTACA CCTGCGCAAA AAATACCAGC TGGAGCGGCG ATAGGTGTAC CGCAAA----
SEQ 25    CAGCATGACG GGCACCGCGA ACAAGGCCGC CCCCGGTGTG CCGTTTTACA CCCCGGCCCA GGAGCCTCCC GCGGGAACGC CAGTCGACGC CAGCACGG--
SEQ 26    ----ATGACG GGCACCGCGA ACAAGGCCGC CCCCGGTGTG CCGTTTTACA CCCCGGCCCA GGAGCCTCCC GCGGGAACGC CAGTCGACGC CAGCACGG--
SEQ 28    GGCTTACGAG ATAATCGACA ACGTTGCGGC TGAAGGGGTT CCATATTACA CACCGGCTCA AGACCCGCCA GCTGGTACGC AGACAAGCGG CTCAACG---
SEQ 29    GGCTTACGAG ATAATCGACA ACGTTGCGGC TGAAGGGGTT CCATATTACA CACCGGCTCA AGACCCGCCA GCTGGTACGC AGACAAGCGG CTCAACG---
SEQ 32    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 34    CCATCACAAA ATCATCATCA ATAAGGAAGC TCCGAATGTT CCTTTCTATA CTCCAGTGCA AGATCCACCA GCAGGAACGT CTTACGATGT TCAGCCTGAA
SEQ 36    -GCACGAGGG ATTATTGACA ACATCGCGGC TGAAGGGGCT CCCTACTACA CGCCTGCTCA AGACYCTCCA GCAGGCACAC AGACCAGCGG CTCAACCA--
SEQ 37    -GCACGAGGG ATTATTGACA ACATCGCGGC TGAAGGGGCT CCCTACTACA CGCCTGCTCA AGACYCTCCA GCAGGCACAC AGACCAGCGG CTCAACCA--
SEQ 39    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 41    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 43    ---------- ---------- ---------- ---------- -------ATGT CCCCACCACG CTTCGAAGCG GCCCCTGCCG ACCCCTCACC GCTCGGC---
SEQ 82    AAACAAGGAG GTTGTTCAGA ATGTCGCTGC CAAAGGAGTG CAATACTTCA ACCCTGAGCA ACTTCCTGCA CCAGGTCTCG GTATAAACGG TCCCAAT---
SEQ 84    ACCGCCTCTC GTCGACTCGA TCGATGCACT CAAGATCAGC AACTTTGTCC CCACTCGAAG TGGCCACCCT CCTCCTGGCT CGGTCCCGGA ATCCATCCTG

          401        411        421        431        441        451        461        471        481        491
          ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
          ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 1     GG------CC AGAAGATCCC CAAGCTCTTC ACGCCCTTGA CCATCCGTGG CGTCACC--- ---------- ---------- -TTCCAGAAC CGCCTTGGTG
SEQ 2     GG------CC AGAAGATCCC CAAGCTCTTC ACGCCCTTGA CCATCCGTGG CGTCACC--- ---------- ---------- -TTCCAGAAC CGCCTTGGT-
SEQ 4     GG------AT CGGCACCTCC CAAGCTCTTC CGGCCGCTTT CGGTGCGGGG TCTGACC--- ---------- ---------- -TTTCACAAT CGCATTGGCG
SEQ 5     GG------AT CGGCACCTCC CAAGCTCTTC CGGCCGCTTT CGGTGCGGGG TCTGACC--- ---------- ---------- -TTTCACAAT CGCATTGGC-
SEQ 7     ---------- ----GACCCC TACGCTCTTC CGGCCCTTAC AAATCCGCAA TGTGACG--- ---------- ---------- -CTCAAGAAC CGCATCATG-
SEQ 9     GG------CA ATGCCGTCCC CAAGCTGTAC ACACCTCTGA CGGTGCGTGG GGTGACC--- ---------- ---------- -TTCCACAAC AGACTTGGC-
SEQ 11    GA------TG AAGTTGCTCC CAAAATTTTT CAACCTTTAA AGATTGGTAA GCTTGCT--- ---------- --------AC ATTACCTAAT AGAATTGGT-
SEQ 13    ---------- ---------- -GCATTATTT CAACCCATAA AGATCAGTGA CTCGATC--- ---------- ---------- -CTCCAGAAC CGCTTCGCC-
SEQ 15    ----------G GCGATGTCCC TACTCTCTTC ACCCCTCTCA AGATCCGTGG TGTTGAG--- ---------- --------AC CCTCCCCAAC CGCCTCGTC-
SEQ 17    ---------- ---------- -AAACTCTCC CAACCCCTCA CCCTCCCCAA TGGCCTT--- ---------- --------AC CCTCCCCAAC CGCCTCGTC-
SEQ 18    ---------- ---------- -AAACTCTCC CAACCCCTCA CCCTCCCCAA TGGCCTT--- ---------- --------AC CCTCCCCAAC CGCCTCGTC-
SEQ 20    ---------- ---CCATCCC AACGCTGTTC AAGCCTCTGA AGATCCGTGG CGTCGAG--- ---------- ---------- -CTCTCCAAC CGCTTTGGC-
SEQ 21    ---------- ---CCATCCC AACGCTGTTC AAGCCTCTGA AGATCCGTGG CGTCGAG--- ---------- ---------- -CTCTCCAAC CGCTTTGGC-
SEQ 23    ---------- ----------C AAAATTATTT ACTCCTCTTA AAATTAGAGG AGTGGAG--- ---------- ---------- -TTCCATAAC AGAATGTTT-
SEQ 25    ---------- ------CTCC GACGCTCTTC AAGCCCCTCC GCATCCGCGA CCTCACC--- ---------- ---------- -ATCAACAAC CGCATCTGG-
SEQ 26    ---------- ------CTCC GACGCTCTTC AAGCCCCTCC GCATCCGCGA CCTCACC--- ---------- ---------- -ATCAACAAC CGCATCTGG-
SEQ 28    ---------- ---------- -AAGCTATTC ACACCCATCA CCATCCGCGG CGTCACA--- ---------- ---------- -TTCCCAAAC CGCCTCTTC-
SEQ 29    ---------- ---------- -AAGCTATTC ACACCCATCA CCATCCGCGG CGTCACA--- ---------- ---------- -TTCCCAAAC CGCCTCTTC-
SEQ 32    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 34    GG-------- ---------- AAGCCTATTC TCTCTTATTA AAATAAGAAA CCTGACT--- ---------- ---------- -CTTCAAAAC CGGATTTTT-
SEQ 36    ---------- ---------- --AGGTTTTC ACACBCATCA CCATCCGAGG CGTCACA--- ---------- ---------- -TTCCCAAAC CGTCTCTTT-
SEQ 37    ---------- ---------- --AGGTTTTC ACACBCATCA CCATCCGAGG CGTCACA--- ---------- ---------- -CTCCAGAAC CGTATTATG-
SEQ 39    ---------- ---------- ---------- -----CCTCA AGATCCGAGG TCTTACC--- ---------- ---------- ---------- ----------
SEQ 41    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 43    ---------- ------ACGC CGCTCAAATA CCCCGTCTCG GGGCGGTCG- ---------- ---------- ---------- -GCGCCCAAC CGGTTCCTC-
SEQ 82    ----------A ATACTCTACC AAAGGTCTTT ACACCCATCA AGATTCGCGG CATGACC--- ---------- ---------- -ATGCCCAAC CGTATCTGG-
SEQ 84    CCAGAGGGTG TCAAAAAACC GGCTTTGTTC CAAACGTTGA CATTGCCCTT TGCTGCACCG GAACAGGCGG GTAAGATGAC CTTCAAGAAC CGCATCATT-

          501        511        521        531        541        551        561        571        581        591
          ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
          ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- --------***
SEQ 1     TAAGTCCGTT TGCCCTTGCT CATATCGACG AAAGCTAATC CCCCGTCAG- ---------- ---------- ---------- -----CTCGC GCCCCTCTGC
SEQ 2     ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- -----CTCGC GCCCCTCTGC
SEQ 4     TGAGTGCAGT CCAGGCAATT ATGCTATCCA TCCTATGCGA GCCCTTGCAT TGGAACAGCC GCTTACAGGG AATGATAATG AGTAGCTATC GCCACTCTGC
SEQ 5     ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- -----CTATC GCCACTCTGC
SEQ 7     ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- -----GTGTC GCCCATGTGC
SEQ 9     ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- -----CTCGC GCCGCTCTGC
SEQ 11    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- -----GTATC TCCAATGTGT
SEQ 13    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- -----GTTTC ACCAATGTGC
SEQ 15    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- -----GTTGC GCCCATGTGC
SEQ 17    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- -----AAAGC CGCCATGGCC
SEQ 18    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- -----AAAGC CGCCATGGCC
SEQ 20    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- -----GTCTC GCCCATGTGC
SEQ 21    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- -----GTCTC GCCCATGTGC
SEQ 23    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- -----GTTTC GCCGATGTGC
SEQ 25    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- -----GTCAG CCCCATGTGC
SEQ 26    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- -----GTCAG CCCCATGTGC
SEQ 28    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- -----CTTGC CCCTCTCTGC
SEQ 29    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- -----CTTGC CCCTCTCTGC
SEQ 32    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 34    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- -----GTCTC CCCAATGTGT
SEQ 36    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- -----CTTGC CCCTCTCTGT
SEQ 37    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- -----CTTGC CCCTCTCTGT
SEQ 39    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- -----TTGAG GGGGCTCTGC
SEQ 41    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 43    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- -----AACGC GGCCATGTCG
SEQ 82    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- -----GTCAG CCCCATGTGC
SEQ 84    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- -----GTCTC TCCCATGTGC
```

```
                601        611        621        631        641        651        661        671        681        691
                --2------- ---------- ---------- ---------- ---------- ---------- ---------- 3--------- ---------- ----------
                *********- ---------- ---------- ---------- ---------- ---------- ********** ********** *--------- ----------
SEQ 1           CAATACTCCG CC-------- ---------- ---CAGGACG GCCACATGAC CGAC------ TACCACATCG CCCATCTGGG TGGGATCGCC CAACGCGGAC
SEQ 2           CAATACTCCG CC-------- ---------- ---CAGGACG GCCACATGAC CGAC------ TACCACATCG CCCATCTGGG TGGGATCGCC CAACGCGGAC
SEQ 4           CAATACTCAG CC-------- ---------- ---GACGATG GACACATGAC TCCC------ TGGCATATGG CACATCTTGG AGGGATTGCC CAGCGAGGGC
SEQ 5           CAATACTCAG CC-------- ---------- ---GACGATG GACACATGAC TCCC------ TGGCATATGG CACATCTTGG AGGGATTGCC CAGCGAGGGC
SEQ 7           ATGTACTCCT GCGAGTCGGA CCCGTCGTCT CCCCACGTCG GCGCCCTAAC AAAC------ TACCACCTGG CGCATCTGGG CCACCTCGCC CTCAAAGGCG
SEQ 9           CAGTACTCCG CA-------- ---------- ---GAAGACG GCCACATGAC AGAC------ TACCACATCG CGCACTTGGG AGGTATTGCC CAGCGCGGCC
SEQ 11          CAATATTCTG CT-------- ---------- GATTATAATT TTGAAGCAAC TCCA------ TACCATTTAA TCCATTATGG TTCATTAGTG AATCGTGGGC
SEQ 13          ATGTATTCAT CG-------- --------TCA CCAACTGACA ATCAAGCCAC TCTG------ TTTCATTTTG TTCATTATGG ATCATTTGCT GTACGTGGAC
SEQ 15          ACCTACTCTG CC-------- ---------- ---GACGATG GCCACATGAC CGAC------ TGGCACCTTG TCCACCTGGG CTCCTTCGCC CTCCGCGGTG
SEQ 17          GAACAAATGG GC-------- ---------- ---TTCGGCA ACCACCTGCC CAAC------ CCCGAACTCG CCGCCGTCTA CGCCACCTGG GCCCGCGGCG
SEQ 18          GAACAAATGG GC-------- ---------- ---TTCGGCA ACCACCTGCC CAAC------ CCCGAACTCG CCGCCGTCTA CGCCACCTGG GCCCGCGGCG
SEQ 20          ACCTACTCAG CC-------- ---------- ---GACGATG GCCACCTGAC CGAC------ TTCCACTTGG TGCACCTGGG CCAGTTCGCC CTGCACGGCA
SEQ 21          ACCTACTCAG CC-------- ---------- ---GACGATG GCCACCTGAC CGAC------ TTCCACTTGG TGCACCTGGG CCAGTTCGCC CTGCACGGCA
SEQ 23          ACTTATTCCG CT-------- ---------- GACCAAGAAG GGCATTTGAC AGAT------ TTTCACCTAG TACATCTTGG AGCGATGGGA ATGCGTGGGC
SEQ 25          CAGTACTCCG CC-------- ---------- ---GACAATG GCCACGCGAC CGAC------ TACCACCTCG TCCACCTGGG CCAGTTCGCC CTGCACGGCG
SEQ 26          CAGTACTCCG CC-------- ---------- ---GACAATG GCCACGCGAC CGAC------ TACCACCTCG TCCACCTGGG CCAGTTCGCC CTGCACGGCG
SEQ 28          CAATACTCCG CC-------- ---------- ---AAAGATG GTTATGCCAC TGAT------ TGGCACTTGA CTCACCTCGG GGGAATAATC CAAAGAGGCC
SEQ 29          CAATACTCCG CC-------- ---------- ---AAAGATG GTTATGCCAC TGAT------ TGGCACTTGA CTCACCTCGG GGGAATAATC CAAAGAGGCC
SEQ 32          ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 34          CAATATTCAG CA-------- ---------- ---AAGGATG GTGTCATGAC CCCC------ TGGCACAAAC AACACCTGGG CAGCTTCGCA GCACGAGGTC
SEQ 36          CAATACTCCG CC-------- ---------- ---AAAGATG GATATGCTAC TGAT------ TGGCACTTGA CTCATCTCGG AGGCATTATC CAACGAGGCC
SEQ 37          CAATACTCCG CC-------- ---------- ---AAAGATG GATATGCTAC TGAT------ TGGCACTTGA CTCATCTCGG AGGCATTATC CAACGAGGCC
SEQ 39          CAGTACTCTG CT-------- ---------- ---CCCGACG GACACTACAC AATG------ TGGCATCACA CCCACATGGG CGGCATCATC CAACGCGGTC
SEQ 41          ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 43          GAGGGCCTGG CG---ACGTT TGACGAGGCG GACCCGTCCA AGCGCGGCAT CCCGACGGAG CAGCTGGTGC AGCTGTACCG GCGCTGGGGC CAGGGCGAGT
SEQ 82          CAATACAGTG CC-------- ---------- ---CGTGACG GCTTTCAGCA GCCT------ TGGCACTTTG CCCACTACGG CGGACTGGCC CAACGTGGCC
SEQ 84          CAGTACTCTG CG-------- ---------- ---AACAATG GTCTTCCTAC TCCG------ TACCACATTG CGCATTTGGG ATCGTTTGCC CTGCACGGTG

                701        711        721        731        741        751        761        771        781        791
                ---------- ---------- ---------- ---------- ---------- ------4---- ---------- ---------- ---------- ----------
                ---------- ---------- ---------- ---------- -------*** ********** *****----- ---------- ---------- ----------
SEQ 1           CCGGCCTGAT GCTGATTGAG GCGACCGCCG TCCAGCCCGA A---GGCCGC ATCACCCCTC AGGATGTCGG TCTGTGGAAG GACTCC---- --------CA
SEQ 2           CCGGCCTGAT GCTGATTGAG GCGACCGCCG TCCAGCCCGA A---GGCCGC ATCACCCCTC AGGATGTCGG TCTGTGGAAG GACTCC---- --------CA
SEQ 4           CAGGATTCTT GATGGTCGAG GCAACAGCAG TCGAACCGGA A---GGCAGG ATCACCCCGC AGGACCTGGG ACTATGGAAA GACTCG---- --------CA
SEQ 5           CAGGATTCTT GATGGTCGAG GCAACAGCAG TCGAACCGGA A---GGCAGG ATCACCCCGC AGGACCTGGG ACTATGGAAA GACTCG---- --------CA
SEQ 7           CAGGCCTCGT CTTCATCGAA GCCACCGCCG TGCAGCCCAA C---GGGCGC ATCTCCCCCA ACGACTCGGG CCTCTGGCAG GACGGCACCA CCTCGGAACA
SEQ 9           CCGGTCTCAT GATGATCGAG GCAACCTCCG TCTCACCTGA A---GGCAGA ATCACGCCGC AGGACGTCGG TTTATGGAAG GACTCG---- --------CA
SEQ 11          CAGGTATCAC CATTGTTGAA AGCACGGCTG TTTCTCCTGA G---GGTGGA TTATCACCTC ATGATTTAGG AATCTGGAAG GATGAA---- --------CA
SEQ 13          CAGCATTAAT CATTTTAGAG AGTATCTTTG TGTCCGAAAA T---TCCGGA TTATCCATTC ATGATTTAGG TCTTTGGAAT GATGAT---- --------CA
SEQ 15          TCCCCCTCAC CATCTTCGAG GCCACCGGCG TCCTCCCCAA C---GGCCGC ATCACCCCCG AGTGCTCTGG TCTCTGGCAG GACTCC---- --------CA
SEQ 17          ACTGGGGCCT GATTCTCACC GGCAACGTCC AAGTCGACCA CGCGCACAAG GGCGACGCCC ACGACATCAG CCCCAACCAC CCCGGCACCA CGCCCGAGCA
SEQ 18          ACTGGGGCCT GATTCTCACC GGCAACGTCC AAGTCGACCA CGCGCACAAG GGCGACGCCC ACGACATCAG CCCCAACCAC CCCGGCACCA CGCCCGAGCA
SEQ 20          CGGCCCTGAC CATTGTCGAG GCCACATCCG TCACGCCCAA C---GGACGC ATCTCGCCCG AGGACAGCGG CCTGTGGCAA GACAGC---- --------CA
SEQ 21          CGGCCCTGAC CATTGTCGAG GCCACATCCG TCACGCCCAA C---GGACGC ATCTCGCCCG AGGACAGCGG CCTGTGGCAA GACAGC---- --------CA
SEQ 23          CTGGCTTTGT AATGGTAGAA GCGACAGCGG TTTCCCCAGA G---GGACGA ATTTCACCTA ATGATTCAGG ATTATGGATG GAGTCG---- --------CA
SEQ 25          CCGCCCTGTC CATGGTCGAG GCCACCGCCG TCGAGGCTCG T---GGCCGC ATCTCCCCCG AGGATGTCGG TTTGTGGCAG GACTCG---- --------CA
SEQ 26          CCGCCCTGTC CATGGTCGAG GCCACCGCCG TCGAGGCTCG T---GGCCGC ATCTCCCCCG AGGATGTCGG TTTGTGGCAG GACTCG---- --------CA
SEQ 28          CCGGATTGTC CATGGTGGAG GCTACCGCTG TACAAAACCA C---GGTCGC ATCACACCTC AGGATGTTGG TCTGTGGGAA GACGGC---- --------CA
SEQ 29          CCGGATTGTC CATGGTGGAG GCTACCGCTG TACAAAACCA C---GGTCGC ATCACACCTC AGGATGTTGG TCTGTGGGAA GACGGC---- --------CA
SEQ 32          ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 34          CGGGTCTCAT TGTCACAGAA GTCAACGCAG TTTCACCAGA G---GGACGA ATCAGTCCTG AGGATGCAGG CATCTACGAT GATGGG---- --------CA
SEQ 36          CGGGACTGTC CATGGTAGAG GCCACCGCTG TTCAAAACCA C---GGTCGC ATCACGCCTC AGGACGTTGG TCTCTGGGAA GATGGA---- --------CA
SEQ 37          CGGGACTGTC CATGGTAGAG GCCACCGCTG TTCAAAACCA C---GGTCGC ATCACGCCTC AGGACGTTGG TCTCTGGGAA GATGGA---- --------CA
SEQ 39          CCGGACTCAC CTGCGTTGAA GCCACAGCCG TGACTCCTCA A---GGTCGC ATCACGCCTG AAGACGTCGG TATCTGGCAA GATTCT---- --------CA
SEQ 41          ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 43          GGGGCCAGAT CCAGACGGGC AACGTCATGA TCGACCCGGA GCACCTCGAG GCCCCGGGCA ACATGGTGGT GCCGCGCGAC GCCGAGCCCT CGGGCGAGCG
SEQ 82          CTGGCCTCAT CATGCTAGAA GCTACCGCAG TTCAAGCACG T---GGCCGT ATCACACCTG AAGATTCTGG CATCTGGCTA GACTCT---- --------CA
SEQ 84          TGGGAAACGT CATGGTCGAA GCATCTGGTG TTGAGCCAGA G---GGGAGG ATCACGCCTC AGGACCTGGG TATTTGGTCG GAACAG---- --------CA

                801        811        821        831        841        851        861        871        881        891
                ---------- ---------- ---------- ---------- -------5-- ---------- ---------- ---------- ---------- ----------6
                ---------- ---------- ---------- ---------- *********- ******---- ---------- ---------- ---******* **********
SEQ 1           GATCGCCCCG ---ATGCGCC GGGTCATCGA CTTCGTGCAC AGCCAGGGC- CAGAAGATCG GCGTG----- ----CAGCTT GCCCATGCCG GCCGGAAAGC
SEQ 2           GATCGCCCCG ---ATGCGCC GGGTCATCGA CTTCGTGCAC AGCCAGGGC- CAGAAGATCG GCGTG----- ----CAGCTT GCCCATGCCG GCCGGAAAGC
SEQ 4           GATTGAGCCA ---TTGAGCC GCGTGATCGA GTTTGTCCAC AGTCAGAAC- CAGCTTATCG GCGTG----- ----CAGATC GCACACGCAG GTCGCAAGGC
SEQ 5           GATTGAGCCA ---TTGAGCC GCGTGATCGA GTTTGTCCAC AGTCAGAAC- CAGCTTATCG GCGTG----- ----CAGATC GCACACGCAG GTCGCAAGGC
SEQ 7           ATTCCTGGGG ---CTGAAGC GGGTCGTCGA GTTCATGCAC GCACAGGGC- GCCAAGGTCG GGATC----- ----CAGATT GCCCACGCCG GCCGCAAGGC
SEQ 9           GATTGCGCCC ---ATGAAGC GCGTCATCGA CTTCGTGCAC TCGCAGTCC- CAGAAGATTG GCGTG----- ----CAGATT GGCCATGGTG GTAGAAAAGC
SEQ 11          AGCAGAGAAA ---TTGAAAC CAATTGTCGA TTACGCTCAT TCTCAAAAG- CAATTAATTG CCATC----- ----CAATTG GGCCATGGTG GTAGAAAAGC
SEQ 13          AGCTCACAGT ---TTACGGA AAATTGTTGA TTTTATTCAT GATCAAGAC- GGAATTTGCT GTATA----- ----CAATTG GCCCACGCCG GCCGCAAGGC
SEQ 15          GATTGCGCCC ---CTCAAGC GCATCGTCGA CTACATCCAC TCCCAGGGC- CAGAAGGCCG GTATC----- ----CAGCTT GCCCACGCCG GCCGCAAGGC
SEQ 17          GACCGTCACG GCCTTCAAGG CCTGGGCGGA CGCCGCGCGC CTGAATGGC- CAGTCCAAAA CGCCTGTGGT CGTGCAGATC AACCACCCTG GTCGCCAGAG
SEQ 18          GACCGTCACG GCCTTCAAGG CCTGGGCGGA CGCCGCGCGC CTGAATGGC- CAGTCCAAAA CGCCTGTGGT CGTGCAGATC AACCACCCTG GTCGCCAGAG
SEQ 20          GATCGCTCCT ---CTGCGCC GCATCGTCGA CTACGTGCAC AGCCAGGGC- CAAAAGATCG CCATC----- ----CAACTG GCTCATGCCG GCCGCAAGGC
SEQ 21          GATCGCTCCT ---CTGCGCC GCATCGTCGA CTACGTGCAC AGCCAGGGC- CAAAAGATCG CCATC----- ----CAACTG GCTCATGCCG GCCGCAAGGC
SEQ 23          AATGAAGCCG ---TTACGAA GAATTGTTGA ATTTGCTCAT TCGCAAAAT- CAAAAAATTG GGATT----- ----CAATTG GCGCATGCTG GTAGAAAGGC
SEQ 25          GATTGCGCCG ---CTGAAGC GCATCGTCGA CTTTATCCAC TCGCAGAAC- CAGGTCGCGG CCATC----- ----CAGCTC GCCCACGCCG GTCGCAAGGC
SEQ 26          GATTGCGCCG ---CTGAAGC GCATCGTCGA CTTTATCCAC TCGCAGAAC- CAGGTCGCGG CCATC----- ----CAGCTC GCCCACGCCG GTCGCAAGGC
SEQ 28          GATCGAGCCT ---CTGAAGC GCATCACCAC TTTCGCGCAC AGTCAGAGC- CAGAAAATTG GTATC----- ----CAGCTG TCGCATGCGG GTCGCAAGGC
SEQ 29          GATCGAGCCT ---CTGAAGC GCATCACCAC TTTCGCGCAC AGTCAGAGC- CAGAAAATTG GTATC----- ----CAGCTG TCGCATGCGG GTCGCAAGGC
SEQ 32          ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 34          GCTTGGACCT ---CTCCGGG ATATTGTGGA CTTTGTACAC AGCCAGGGC- GCCAAGATTG CTATT----- ----CAGATA GGTCATGCTG GGAGAAAAGC
SEQ 36          AATCGAGCCC T--TTGAAGC GCATCACTAC TTTTGCCCAC AGCCAAAGCW CAGAAGATTG GTAT------ ---TCAGCTC TCGCACGCTG GTCGTAAGGC
SEQ 37          AATCGAGCCC ---TTGAAGC GCATCACTAC TTTTGCCCAC AGCCAAAGC- CAGAAGATTG GTAT------ ---TCAGCTC TCGCACGCTG GTCGTAAGGC
SEQ 39          GATCGAGCCT C--TTGCCAA GGTCGTC-GA GTTTGCCCAC TCCCAGAAC- CAGAAGATCA TGATT----- ----CAGTTG GCGCATGCGG GCCGCAAAGC
SEQ 41          ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 43          CTTCGACATG TTTTCCAAGC TCGCCGCCGC CGCCAAGGAG CACGGCAGC- CTC-ATCGTC GCG------- ----CAGGTC GGACACCCCG GTCGCCAGGC
SEQ 82          TGTTGAGGGA ---CTGCGAA AGCACGTCGA GTTTGCCCAT GCCAACAAC- TCTCTTATCG GTATC----- ----CAGATT GGCCATGCTG GTCGCAAGGC
SEQ 84          TCGGGATGCA ---CACAAGG CGCTGGTGTC GGTGCTCAAG TCCTTCACG- GATGGTCTGG GTGTA----- ----GGGCTG CAACTGGCGC ATGCGGGAAG
```

```
          901        911        921        931        941        951        961        971        981        991
          ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
          ********** **********_ ---------- ---------- ---------- ---------- ---------- ---------- ---------** **********
SEQ 1     CACCACCGTT GCGCCCTGGA TCTCA----- ---------- ---------- ---------- -TTCTCGGCC ATCGCGACGG AGAAGGTCGG CGGATGGCCG
SEQ 2     CACCACCGTT GCGCCCTGGA TCTCA----- ---------- ---------- ---------- -TTCTCGGCC ATCGCGACGG AGAAGGTCGG CGGATGGCCG
SEQ 4     CAGCACCGTC GCGCCATGGC TCTCG----- ---------- ---------- ---------- -GCCAACGAT ACCGCCTCCG AGAAGATGGG CGGCTGGCCA
SEQ 5     CAGCACCGTC GCGCCATGGC TCTCG----- ---------- ---------- ---------- -GCCAACGAT ACCGCCTCCG AGAAGATGGG CGGCTGGCCA
SEQ 7     GAGTGCCGTT GCGCCGTGGC TGGCG----- ---------- ------GCGC AGGCGGGCAA GTCGAGTCTG AAGGCGGATG AGAGCGTTGG CGGGTGGCCC
SEQ 9     TTCGAACATC GCCCCCTGGC TCATG----- ---------- ---------- --------AA CAAGGGCATC GTCGCGACGG AGAAGGTCGG TGGCTGGCCG
SEQ 11    TTCTGGTCAG CCCTTATTTT TGCAC----- ---------- ---------- ---------- -TTGGAACAA GTTGCAGATA AATCTGTCAA TGGGTTTGCC
SEQ 13    TGTTGAAGGG GTACCATTCC AACAA----- ---------- ---------- ---------- ---------- ---------- --ATACAACA TGGTTGGCAA
SEQ 15    CTCCACCAAG GCCCCCTGGC ACTAC----- ---------- ---------- --CAGCGCGG CAAGAGCGAG CTTGCCGGCC CCGAGCAGGG TGGCTGGCCC
SEQ 17    TCCGATGGGC GCGGGCACGC GGGGA----- ---------- ------CTGT GGGAGAAGGC GGTGGCGCCC TCGCCGGTGC CGTTGGTGTT GGGAGAGGCG
SEQ 18    TCCGATGGGC GCGGGCACGC GGGGA----- ---------- ------CTGT GGGAGAAGGC GGTGGCGCCC TCGCCGGTGC CGTTGGTGTT GGGAGAGGCG
SEQ 20    CAGCACAAAG GCCCCCTGGC ACGACTCCTT CACCCCCAGC GGCGAGTATA AGCCGAGAGA GGGCTTACAG GTCGTCGGAC CCGAGTATGG CGGCTGGCCT
SEQ 21    CAGCACAAAG GCCCCCTGGC ACGACTCCTT CACCCCCAGC GGCGAGTATA AGCCGAGAGA GGGCTTACAG GTCGTCGGAC CCGAGTATGG CGGCTGGCCT
SEQ 23    TAGCACCACT GCTCCTTATC GAGGA----- ---------- ---------- ---------- ----TACACA GTTGCGACTG AAGCTCAAGG TGGGTGGGAG
SEQ 25    TAGCACCCTG GCACCGTGGA TCACC----- ---------- ---------- --GAGGCTCG CGGCAAGGCG CTGGCTCAGG AGAGCGAGAA CGGCTGGCCC
SEQ 26    TAGCACCCTG GCACCGTGGA TCACC----- ---------- ---------- --GAGGCTCG CGGCAAGGCG CTGGCTCAGG AGAGCGAGAA CGGCTGGCCC
SEQ 28    CAGTTGCGTA TCTCCCTGGC TAAGC----- ---------- ---------- ---------- -GTAAATGCT GTCGCGGCGG AAGAAGTGGG TGGCTGGCCA
SEQ 29    CAGTTGCGTA TCTCCCTGGC TAAGC----- ---------- ---------- ---------- -GTAAATGCT GTCGCGGCGG AAGAAGTGGG TGGCTGGCCA
SEQ 32    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 34    GAGCACAGTC GTACCGTGGC TGGAC----- ---------- ---------- ---------- -CGCAAGAAC ACTGCTTTTA ---------- ----------
SEQ 36    TAGTTGTGTA TCTCCGTGGT TGAGC----- ---------- ---------- ---------- -ATCAACGCT GTTGCCGCTA AGGAAGTCGG TGGCTGGCCA
SEQ 37    TAGTTGTGTA TCTCCGTGGT TGAGC----- ---------- ---------- ---------- -ATCAACGCT GTTGCCGCTA AGGAAGTCGG TGGCTGGCCA
SEQ 39    GAGCACTGTG GCACCATGGT TAAGC----- ---------- ---------- ---------- -GGCGGCGAT GTTGCTGGTG AGGACGTCAA CGGATGGCCA
SEQ 41    ---------- ---------- ---------- ---------- ---------- ------GACT GCCGAGTAAA CGCGCCGGCA AGGAGGCGGG AGGATGGCCG
SEQ 43    CCGCGGCAGC GTCCAGCAGC ACCCC----- ---------- ---------- --ATTAGCGC CAGCGACGTG CAGCTTAAGC AGGAGATG-- ----------
SEQ 82    CTCCTGCGTT GCTCCTTGGT TAGAC----- ---------- ---------- ---------- -GCCGGACTT GCCGCTGAAA AGGCCGCTGG TGGATGGCCC
SEQ 84    GAAGGCCTCG GACTGGTCAC CTTTC----- ---------- ------TACC GCGGAGAAAA GAAGCAAAAG TTTGTGACGC AGGAGGAAGG TGGCTGGCCG

          1001       1011       1021       1031       1041       1051       1061       1071       1081       1091
          ---------- ---------- -7-------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
          ********** ********** ********-- ---------- ---------- ---------- -********* ********** *****----- ----------
SEQ 1     GACCCGCGTC AAAGGGCCCG GCGATATC-- ---------- ---------- ---------- -CCCTTTGCG GAGCCCTTCG CCAAGCCCAA GGCCATGACG
SEQ 2     GAC-CGCGTC AAAGGGCCCG GCGATATC-- ---------- ---------- ---------- -CCCTTTGCG GAGCCCTTCG CCAAGCCCAA GGCCATGACG
SEQ 4     GGC-CGCGTC AAAGGCCCGA CAAATGTG-- ---------- ---------- ---------- -CCCTTCACC GTTAAGAACC CTGTGCCGAA GGAGATGACC
SEQ 5     GGC-CGCGTC AAAGGCCCGA CAAATGTG-- ---------- ---------- ---------- -CCCTTCACC GTTAAGAACC CTGTGCCGAA GGAGATGACC
SEQ 7     GCG-GATGTG GTGGGTCCGT CGGGCGGG-- ---------- ---GAGGAGC ATATCTTTAG TCCCGAGGAG GATGCGTATT GGGTGCCGCG GGCGCTGAGC
SEQ 9     GAT-CGTGTG ATCGGCCCGT CCACCGTG-- ---------- ---------- ---------- -CCCTTCCAC GAGACTTTCC CCACCCCCAA GGCCATGACC
SEQ 11    GAC-AAAGCA GTTGCTCCTT CTGCATTG-- ---------- ---GCATTC- ---------- -AGACCAAAT GGTAATTTAC CTGTTCCTAA TGAGTTGACC
SEQ 13    GAA-CATTGT GTGGGGCCAT CTACTGAG-- ---------- ---------- ---------- -CCATTTAGT GATTCACACA ATACACCACG AGAATTGACT
SEQ 15    GAG-AACGTC TGGGCCCCCA GCGCCATC-- ---------- ---------- ---------- --------AG CTACAACGAG GAGACCTTCC CCTTCCCCAA GGAGATGACC
SEQ 17    TTT-GTGCCT CGCTTGTTGT CGAAAGTG-- ---------- ---------- ---------- ---------- ---CTTTTCG GCACGCCGCG GGAGCTGACG
SEQ 18    TTT-GTGCCT CGCTTGTTGT CGAAAGTG-- ---------- ---------- ---------- ---------- ---CTTTTCG GCACGCCGCG GGAGCTGACG
SEQ 20    GAT-GACGTC TGGGCCCCGA GCGCCATC-- ---------- ---------- ---------- -CCGTTCTCG GAGGACTTTC CGAACCCCAA GGAGATGACC
SEQ 21    GAT-GACGTC TGGGCCCCGA GCGCCATC-- ---------- ---------- ---------- -CCGTTCTCG GAGGACTTTC CGAACCCCAA GGAGATGACC
SEQ 23    AAT-GATGTT TATGGACCAA ATGAAGAC-- ---------- ---------- ---------- -AGGTGGGAC GAAAACCACG CTCAACCTCA TAAGTTAACT
SEQ 25    GAC-GACGTT GTGGCTCCCA GCGCGATT-- ---------- ---------- ---------- -CCTTACACC AAGGACTGGG CCACACCGCG TGAGTTGACT
SEQ 26    GAC-GACGTT GTGGCTCCCA GCGCGATT-- ---------- ---------- ---------- -CCTTACACC AAGGACTGGG CCACACCGCG TGAGTTGACT
SEQ 28    GAC-AATATC GTTGCTCCCT CGGCCATC-- ---------- ---------- --------GC ACAAGAAAAT GGTGTGAACC CAGTTCCCAA GGCTTTCACG
SEQ 29    GAC-AATATC GTTGCTCCCT CGGCCATC-- ---------- ---------- --------GC ACAAGAAAAT GGTGTGAACC CAGTTCCCAA GGCTTTCACG
SEQ 32    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 34    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 36    GAC-AACATT GTTGCTCCTT CTGCCATC-- ---------- ---------- --------GC ACAAGAAGCT GGCGTGAACC CTGTTCCCAA GGCCTTCACC
SEQ 37    GAC-AACATT GTTGCTCCTT CTGCCATC-- ---------- ---------- --------GC ACAAGAAGCT GGCGTGAACC CTGTTCCCAA GGCCTTCACC
SEQ 39    CAG-GATGTC TGGGCGCCCA GTGCGATT-- ---------- ---------- ---------- -CCATGGAAC GAGAAGCACG CTGTCCCAAA GGAGATGTCG
SEQ 41    GAG-GATGTT GTGGGTCCGT CGGGTGGGGA GGACTTTACG TGGGATGAGA GGTCCTCGAG CGACCCTAGT GGAGGCTACT ATGCGCCGAG AGAGTTGTCG
SEQ 43    ---------- ---------- ---------- ---------- ---------- ---------- ----TTTGGG TCAAAGTTTG GCGTGCCCAG GCCCGCTACC
SEQ 82    GAT-GACGTT GTCGGACCTA GCAACGAG-- ---------- ---------- ---------- -CGTTTTGCT CCTGGCTACC CTACCCCCCG TGCTATTACT
SEQ 84    GAT-CGTGTC GTCGCTCCTT CGGCCATC-- ---------- ---------- ---------- -GCATATGCG CAAGGTCACG TTACCCCTCG AGCTCTCACG

          1101       1111       1121       1131       1141       1151       1161       1171       1181       1191
          ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
          ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 1     CTGGATGA-G ATCGAGCAGT TCAAGAAGGA CTGGGTGGCG GCCACGAAGC GCGCCATCGC CG---CCGGT GCGGACTTTG TCGAGATTCA CAATGCGCAT
SEQ 2     CTGGATGA-G ATCGAGCAGT TCAAGAAGGA CTGGGTGGCG GCCACGAAGC GCGCCATCGC CG---CCGGT GCGGACTTTG TCGAGATTCA CAATGCGCAT
SEQ 4     AAGCAGGA-T ATCGAGGATC TGAAGACCGC CTGGGTGGCC GCTGTCAAAC GGGCTGTTAA GG---CCGGA GCCGACTTTA TCGAGATCCA CAATGCGCAT
SEQ 5     AAGCAGGA-T ATCGAGGATC TGAAGACCGC CTGGGTGGCC GCTGTCAAAC GGGCTGTTAA GG---CCGGA GCCGACTTTA TCGAGATCCA CAATGCGCAT
SEQ 7     ACGGCCGA-G GTCCGTCAGG TGGTGGCGGC GTTTGCGAAG AGCGCGCGGC TAGCGGTGCA GG---CTGGG GTGGATGTTA TCGAGATCCA TGGGGCGCAT
SEQ 9     AAGGACGA-C ATCGAGCAGT TCAAGCGCGA CTGGTTTGAT GCGTGCAAGC GGGCCATTGC CG---CTGGC GCGGACTTCA TCGAGATCCA CAATGCCCAC
SEQ 11    AAAGATGA-A ATCAAACGTG TTGTTAAGGA TTTTGGTGCT GCTGCTAGAA GAGCTGTTGA AATCAGTGGC TTTGATGCAG TTGAGATTCA TGGTGCTCAT
SEQ 13    GTTAATGA-A ATAAATTCAA TTGTGGAAGA CTTTGCCAAT GCAGCTTGGC GGGCTGTGGA AATCTCAAAA TTCGATGCCA TTGAAATACA TTGTGCTAAT
SEQ 15    GTCGAGCA-G ATCCACGAGC TCGTCGAGGC CTGGAAGGCG TCTGCCCAGC GTGCCCTCAA GG---CCGGC TTCGACCTCA TTGAGATCCA CGCCGCCCAC
SEQ 17    GTTGCGGA-G ATCAAGGATA TCGTGCAAAA GTTTGCGGTG ACGGCGAGGA TCACGGCCGA GG---CCGGG TTCAATGGCG TGGAGATCCA TGCGGCGCAT
SEQ 18    GTTGCGGA-G ATCAAGGATA TCGTGCAAAA GTTTGCGGTG ACGGCGAGGA TCACGGCCGA GG---CCGGG TTCAATGGCG TGGAGATCCA TGCGGCGCAT
SEQ 20    GTTGAGGA-G ATTGAGGGAC TCGTCACCAG CTTTGTGGAC GCTGCCAAGC GTGCCATCGA GG---CCGGC GTCGACATTA TTGAGATTCA CGGCGCTCAC
SEQ 21    GTTGAGGA-G ATTGAGGGAC TCGTCACCAG CTTTGTGGAC GCTGCCAAGC GTGCCATCGA GG---CCGGC GTCGACATTA TTGAGATTCA CGGCGCTCAC
SEQ 23    GAAAAGCA-A TATGATGAAT TAGTGGATAA GTTTGTTGTT GCTGCGAAGC GTGCAGTTGA AA---TAGGT TTTGATGTAA TTGAAATTCA TGGCGCTCAT
SEQ 25    ACCGAGGRRG TCGAGGGTCT GGGTGAAGAA GTTCGCCGAG TCGGCCAAGA GGTCAAATCG A---GCTGGT TTTGACGTCA TTGAGATCCA CGCCGCTCA-
SEQ 26    ACCGAGGR-G TCGAGGGTCT GGGTGAAGAA GTTCGCCGAG TCGGCCAAGA GGTCAAATCG AG---CTGGT TTTGACGTCA TTGAGATCCA CGCCGCT---
SEQ 28    AAGGAGGA-T ATAGAGCAAC TCAAGAGCGA CTACGTGGAA GCGGCAAAAC GAGCCATCCA TG---CTGGT TTCGATGTTA TCGAAATTCA TGCAGCTCAT
SEQ 29    AAGGAGGA-T ATAGAGCAAC TCAAGAGCGA CTACGTGGAA GCGGCAAAAC GAGCCATCCA TG---CTGGT TTCGATGTTA TCGAAATTCA TGCAGCTCAT
SEQ 32    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 34    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 36    AAGGAGGA-T ATCGAGGAAC TCAAGAATGA CTTTCTGGCT GCAGCMAAAC GAGCCAWCCG CGC---TGGT TTTGATGTCA TCGAGATCCA TGCAGCTCAT
SEQ 37    AAGGAGGA-T ATCGAGGAAC TCAAGAATGA CTTTCTGGCT GCAGCMAAAC GAGCCAWCCG CGC---TGGT TTTGATGTCA TCGAGATCCA TGCAGCTCAT
SEQ 39    TTGGATGA-T ATCGAGGCTT TCAAGAAGGC GTTTGGAGAG GCGGTCAAGC GGGCATTGAA GGC---TGGA TTTGATGTTA TCGAGATTCA CAATGCTCAC
SEQ 41    GTCAGAGA-G ATCAAGGAGA TGGTCCAAGA CTGGGCGACA GCAGCGAAAA GGGCGGTGAA AGC---GGGC GTGGATGTAA TCGAAATCCA CGGCGCGCAT
SEQ 43    AAGGAGGA-T ATTAAGGCGG TGATTGAGGG TTTTGCCCAC ACGGCCGAGT ACCTTGAAAA GGC---CGGT TTCGACGGTA TCGAATTGCA CGCCGCCCAC
SEQ 82    CTTGAAGA-G ATTGAACAGT TGAAGGAGGA CTTTGTTTCC GGTGTTCGTC GAGCGGTTGA AG---CAGGA TTTGACACTA TCGACTTCCA TTTCGCTCAC
SEQ 84    ACCGAGGA-C ATCAACAAGT TGCAAGACAA ATTCGTTCAG TCGGCACGAT GGGCGTTTGA AG---CTGGG TATGACTACG TCGAACTTCA CAGCGCTCAC
```

```
          1201       1211       1221       1231       1241       1251       1261       1271       1281       1291
          ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
          ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 1     GGATACCTGC TGTCGTCATT CCTCTCGCCG GCCGCCAAC- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 2     GGATACCTGC TGTCGTCATT CCTCTCGCCG GCCGCCAAC- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 4     GGCTATCTTC TGATGTCGTT CCTCTCCCCT GCGGTCAAC- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 5     GGCTATCTTC TGATGTCGTT CCTCTCCCCT GCGGTCAAC- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 7     GGCTATCTCA TCAACGAGTT CCTGAGCCCG GTCACGAAT- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 9     GGGTATCTTC TCTCGTCTTT CCTATCACCG TCTTCCAAC- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 11    GGTTATTTGA TTAATGAGTT CTATAGTCCT ATTTCAAAC- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 13    GGATGTTTAA TACACCAATT TTTAAGTAAA TTGACAAAC- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 15    GGCTACCTCA TTTCCGAGTT CTTGAGCCCC ATCTCCAAC- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 17    GGATACCTGT TGGCGCAGTT CTTGAGCAAG AAGACAAAC- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 18    GGATACCTGT TGGCGCAGTT CTTGAGCAAG AAGACAAAC- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 20    GGTTACCTGA TCACCGAGTT CCTTTCGCCG CTATCAAACG TAAGTGGAGA TACTTTGTGT GGGGCTGTGC GCATACTCCC TCGGGTGTGA CTTCTATTAA
SEQ 21    GGTTACCTGA TCACCGAGTT CCTTTCGCCG CTATCAAAC- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 23    GGTTATCTTA TATCGTCAAC AGTTAGTCCT GCCACTAAT- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 25    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 26    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 28    GGATATCTAC TGCATCAATT CTTGAGTCCG GTAAGCAAT- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 29    GGATATCTAC TGCATCAATT CTTGAGTCCG GTAAGCAAT- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 32    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 34    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 36    GGATACKTGC TTCACCAGTT CTTGAGTCCA GTCAGTAAC- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 37    GGATACKTGC TTCACCAGTT CTTGAGTCCA GTCAGTAAC- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 39    GGATACCTCC TCCACGAATT CATCTGCCTG AGAGCAACA- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 41    GGGTACCTCA TCCACGAATT CCTCTCACCC ATTACCAAC- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 43    GGTTACCTGC TGGCCCAATT CCTGTCCGAA ACAACCAAC- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 82    GGTTATCTTG TTTCCAGCTT CCTGTCCCCT GCCACCAAC- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 84    GGATACCTGA TGCACTCGTT CCTCAGCCCG TTGACCAAT- ---------- ---------- ---------- ---------- ---------- ----------

          1301       1311       1321       1331       1341       1351       1361       1371       1381       1391
          ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
          ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 1     ---------- ---------- --AACCGCAC GGACCAGTAC GGCGGGTCGT TCGAGAACCG CATCCGGCTG TCTCTCGAGA TTGCGCAGTT GACTCGGGAC
SEQ 2     ---------- ---------- --AACCGCAC GGACCAGTAC GGCGGGTCGT TCGAGAACCG CATCCGGCTG TCTCTCGAGA TTGCGCAGTT GACTCGGGAC
SEQ 4     ---------- ---------- --ACGAGAAC AGACGAGTAC GGAGGCAGTT TTGAGAATCG CATCCGGCTC AGTCTGGAGA TCGCCAAGCT CACCCGCGAA
SEQ 5     ---------- ---------- --ACGAGAAC AGACGAGTAC GGAGGCAGTT TTGAGAATCG CATCCGGCTC AGTCTGGAGA TCGCCAAGCT CACCCGCGAA
SEQ 7     ---------- ---------- --AAGCGGAC GGATGCGTAC GGCGGGAGCT TTGAGAACCG GACCCGGATC GTGCGCGAGG TTGCGGCGGC TATTCGTGCG
SEQ 9     ---------- ---------- --ACGCGCAC CGACGAGTAC GGCGGCTCCT TTGAGAACCG CATCCGGCTC TCTCTCGAAA TCGCCCAGGT CACCCGTGAC
SEQ 11    ---------- ---------- --AAGAGAAC AGATGAATAC GGTGGCAGTT TTGAAAATAG AACCAGATTT TTAAAGGAAG TTATCGATAG TGTTAAATCA
SEQ 13    ---------- ---------- --AAGAGAGC TGACCAATAC GGGGGCTCAT TTGAAAACAG AGTTAGATTT CTTTTACAAA TAATTGAGAA TATAAAACGA
SEQ 15    ---------- ---------- --CAGCGTAC CGACCAGTAC GGTGGCTCCT TCGAGAACCG CACCCGCGTT CTCCGCGAGA TCATCTCGGC CGTCCGCTCC
SEQ 17    ---------- ---------- --AGGCGCGG GGATGAGTAT GGCGGGTCGG CTGAGAACAG GGCGAGGATT GTTGGGGAGA TTATTAAGGA GTGCAGGAGG
SEQ 18    ---------- ---------- --AGGCGCGG GGATGAGTAT GGCGGGTCGG CTGAGAACAG GGCGAGGATT GTTGGGGAGA TTATTAAGGA GTGCAGGAGG
SEQ 20    CATTTTATTT CCTGGCACGC AGAAACGGAC AGACAAGTAC GGCGGCAGCT TTGAGAACCG CACCCGGGTC CTGATCGATA TTATCAAGGC CGTCCGGGCA
SEQ 21    ---------- ---------- --AAACGGAC AGACAAGTAC GGCGGCAGCT TTGAGAACCG CACCCGGGTC CTGATCGATA TTATCAAGGC CGTCCGGGCA
SEQ 23    ---------- ---------- --GACCGCAA TGACAAGTAT GGTGGGACAT TTGAGAAACG TATTTTGTTT CCTATGGAAG TTGTCCATTC TGTTCGTAAA
SEQ 25    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 26    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 28    ---------- ---------- --CAAAGAAC CGACGAGTAT GG-------- ---------- ---------- ---------- ---------- ----------
SEQ 29    ---------- ---------- --CAAAGAAC CGACGAGTAT GG-------- ---------- ---------- ---------- ---------- ----------
SEQ 32    ---------- ---------- -------AAC CGACGAGTAT GGTGGCAGTT TCGAGAACCG TATCAGAGTT GTCTTGGAAA TCCTTGACCT CATCCGCGCT
SEQ 34    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 36    ---------- ---------- --CAAAGAAC CGATGAGTAT GGTGGCAGCT TCGAGAACCG TATCAGAGTC GTCTTGGAGA TCATTG---- ----------
SEQ 37    ---------- ---------- --CAAAGAAC CGATGAGTAT GGTGGCAGCT TCGAGAACCG TATCAGAGTC GTCTTGGAGA TCATTG---- ----------
SEQ 39    ---------- ---------- --CCAGGACC GACAAGTACG GGCGGAAGCT GGGAAAACCG CACTCGTCTG ACAATGGAAA GTCGTCGACC TTGTCCGCAG
SEQ 41    ---------- ---------- --CGCCGGAC AGATTCTTAC GGCGGTTCTT TCGAAAACCG TACCCGTCTA CTCATTGAAA TCGTAACAGC CGTCCGAGCC
SEQ 43    ---------- ---------- --CAGCGCAC CGACGAGTAC GGCGGCAGCC TCGAAAACCG CATGCGGCTA ATCCTCGAGG TCACGGCCGA GGTCCGCAGG
SEQ 82    ---------- ---------- --AAGCGTAC CGACAAGTAC GGAGGTAGCT TCGAGAACAG AGTGCGCCTT GCTCTCGAGA TTGTCGAGGC TGCACGAGCT
SEQ 84    ---------- ---------- --CAGCGTAC CGACGAGTAC GGCGGTAGCC TGGAGAACCG CGCTCGATTT CTGCTCAACG TTGCCCGTCG AATCCGCCAA

          1401       1411       1421       1431       1441       1451       1461       1471       1481       1491
          ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---8------ ---------- ----------
          ---------- ---------- ---------- ---------- ---------- ---------- ----------* ********** *--------- ----------
SEQ 1     GCCGTCGGCC CTCATGTGCC C--------- ---------- -----GTTTT CCTGCGCATT TCGGCCTCGG ACTGGTGCGA GGAGACCCTG CCGGA-----
SEQ 2     GCCGTCGGCC CTCATGTGCC C--------- ---------- -----GTTTT CCTGCGCATT TCGGCCTCGG ACTGGTGCGA GGAGACCCTG CCGGA-----
SEQ 4     AATGTGCCCA AGGATATGCC T--------- ---------- -----GTCTT CCTGCGGGTC TCCGCCACCG ATTGGCTGGA GGAGGTGCAG CCGAA-----
SEQ 5     AATGTGCCCA AGGATATGCC T--------- ---------- -----GTCTT CCTGCGGGTC TCCGCCACCG ATTGGCTGGA GGAGGTGCAG CCGAA-----
SEQ 7     GTGATTCCCG AGGGGATGCC C--------- ---------- -----CTGTT TCTGCGTATC AGCGCCACGG AGTGGTTGGA GGGTCAGCCG GTGGC-----
SEQ 9     GCCGTCGGCC CCAACGTTCC T--------- ---------- -----GTTTT TCTCCGTGTC TCCGCGACGG ACTGGATCGA GGAGACCCTC CCCGA-----
SEQ 11    AGTATTCCAA ACGATGTTCC A--------- ---------- -----GTGTT TTTGAGAATC TCTGCTGCTG AAAATAGTCC TGATCCA--- ----------
SEQ 13    AAGATAGAAA CA------CC G--------- ---------- -----ATTTT CTTAAAGTTT CCAATGTCAG ATAATTGTAG TGATCCG--- ----------
SEQ 15    GTCATCCCCG AGGACATGCC C--------- ---------- -----CTCTT CGTCCGTGTC TCCGCCACCG AGTGGATGGA GTACACC--- ----------
SEQ 17    CAGGTGACTG AGGCGGTGGG TGAAGAGGAG GCGAAGAAGT TTGTGGTGGG AATCAAGCTG AACAGTGCGG ATTGGCAGGC GGGACGCGAT GGA------A
SEQ 18    CAGGTGACTG AGGCGGTGGG TGAAGAGGAG GCGAAGAAGT TTGTGGTGGG AATCAAGCTG AACAGTGCGG ATTGGCAGGC GGGACGCGAT GGAAAG----
SEQ 20    GTGATTCCCG AGGAGATGCC A--------- ---------- -----CTCTT CGTCCGAATC TCCGCGACCG AATGGATGGA GTACGCCGGC ----------
SEQ 21    GTGATTCCCG AGGAGATGCC A--------- ---------- -----CTCTT CGTCCGAATC TCCGCGACCG AATGGATGGA GTACGCCGGC ----------
SEQ 23    GCAATTCCAG ATAGTATGCC C--------- ---------- -----TTGTT TTATAGAGTA ACGGCTACAG ATTGGTTGCC CAAAGGACAA ----------
SEQ 25    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 26    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 28    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 29    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 32    GCCATCCCCG AAACTACACC T--------- ---------- -----GTCCT CGTTCGTGTC AGTGCAACTG ATTGGTTCGA GTTTGACTCT CAATTCAAAG
SEQ 34    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 36    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 37    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 39    CATT------ ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 41    GCGATGCCCT CCAGCATGCC T--------- ---------- -----CTCTT CCTCCGCCTC TCCTCTACAG AATGGATGGA AGATACCGAC ATCGGC----
SEQ 43    CGGACGAGCA AGAATTTCAT C--------- ---------- -----CTCGG CATCAAAATT AACAGCGTCG AGTTCCAGGA GAAG------ ----------
SEQ 82    GTTATGCCTG AGGACATGCC C--------- ---------- -----TTGTT CACTCGCATC AGTGGAACTG ACTGGCTGGA GAACAACCCT GAG-------
SEQ 84    GAATTCCCCA ACAAGGGT-- ---------- ---------- -----CTCTG GGTGCGCGTC AGCTCCACCG ACTGGGCCGA CCAAGCGCAC CAA-------
```

```
          1501       1511       1521       1531       1541       1551       1561       1571       1581       1591
          ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---9------ ----------
          ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----****** ********** *****#####
SEQ 1     ---------- GCAGAGCTGG AAGTCGGAGG ATACCGTGCG GTTCGCGCAG GAGCTGGTCA AGCAGGGCGC CGTTGATCTG ATCGATATCA GCAGCGGTGG
SEQ 2     ---------- GCAGAGCTGG AAGTCGGAGG ATACCGTGCG GTTCGCGCAG GAGCTGGTCA AGCAGGGCGC CGTTGATCTG ATCGATATCA GCAGCGGTGG
SEQ 4     -------CAA GCCCAGCTGG CGAGGCGTGG ACACTGTCCG ATTTGCGAAG ATCCTGGCAG AAACGGGTTA CGTTGACGTG CTTGACGTGA GCAGTGGCGG
SEQ 5     -------CAA GCCCAGCTGG CGAGGCGTGG ACACTGTCCG ATTTGCGAAG ATCCTGGCAG AAACGGGTTA CGTTGACGTG CTTGACGTGA GCAGTGGCGG
SEQ 7     -CGCGGAGTC GGGCAGCTGG GATAT---GC AGAGCTCGCT GGAGCTGGTC AAGAAGCTGC CCGAATGGGG CATTGACCTG GTGGATGTCA GCTCCGCCGC
SEQ 9     ---------- GGAATCGTGG AAGCTCTCTG ACTCCGTCCG CTTCGCCGAA GCCCTCGCTG CCCAGGGCGC TATTGACCTG ATCGACGTCT CTTCCGGCGG
SEQ 11    ---------- -GAAGCTTGG ACTATTGAAG ATTCCAAAA- --AATTAGCT GACATTTTAG TAGAAAAGGG TATTGCTTTG GTTGATGTTT CATCTGGTGG
SEQ 13    ---------- -GAAGCGTGG TCTACGGAAG ATGCATTGA- --AGTTGGCC GATCTTGTTA TTGATTTAGG AGTAAAGGTG ATCGACGTTA CATCAGGTGG
SEQ 15    -----GGCCA GCCCTCGTGG GACCTCCAGC AGACCATTG- --AGCTCGCC AAGATCCTCC CCGACCTCGG CGTCGACCTC CTCGACGTCT CTTCCGGCGG
SEQ 17    AGGAGGAGGA GGAGACGGAT ACGGCGGAGG AGGTGTTGA- --AGCAGATT GAGCTTTTTG AGCAGTGGGG GATCGACTTT GTCGAGGTTA GCGGTGGCAG
SEQ 18    --GAGGAGGA GGAGACGGAT ACGGCGGAGG AGGTGTTGA- --AGCAGATT GAGCTTTTTG AGCAGTGGGG GATCGACTTT GTCGAGGTTA GCGGTGGCAG
SEQ 20    --------GA GCCTAGCTGG GACCTCGAGC AGAGCACAC- --AGCTTGCC AAGCTCCTCC CGGACCTGGG TGTCGACCTG CTCGACGTCA GCTCGGGCGG
SEQ 21    --------GA GCCTAGCTGG GACCTCGAGC AGAGCACAC- --AGCTTGCC AAGCTCCTCC CGGACCTGGG TGTCGACCTG CTCGACGTCA GCTCGGGCGG
SEQ 23    ---------- ----GGATGG GAGATAGAAG ATACAGTTG- --CATTAGCA GCGAGGCTTC GCGATGGTGG TGTTGACTTG ATAGATGTTA GCTCTGGTGG
SEQ 25    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 26    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 28    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 29    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 32    ACGAGTTTCC TGAAAGCTGG ACAGTCGAGC AGACTT---G TCAACTCGCG CGTATCTTGC CCAAGCATGG AGTAGACTTG GTGGACGTCA GCTCAGGCGG
SEQ 34    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 36    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 37    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 39    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 41    --AAGAAGTT CGGAAGCTGG GATGTCGAAA GCACGATCA- --AGATCTCC AAAATCCTGG CCGACTTGGG CGTTGATCTC CTCGACGTGT CTTCCGGTGG
SEQ 43    ---------- -GGTTTCAAG CCA---GAGG AGGCGGTGC- --AGTTGTGC GAGGCCCTCG AGGCCGCGGG CATGGATTTT GTCGAGACGA GCGGCGGCAC
SEQ 82    --TACGAGGG AGAGACCTGG ACTCTTGAGC AGAGCATCA- --AGCTTGCA CACCAGTTAG CAGACCGTGG TGTCGATGTT TTGGATGTTT CCAGTGGTGG
SEQ 84    --------GC CGACTCTTGG ACCGTTGACC AGACGGTTG- --AACTCGCC AAGATGCTCC AAGAGGCTCG AGTCGACCTG CTAGACGTCA GCTCCGGCGG

          1601       1611       1621       1631       1641       1651       1661       1671       1681       1691
          ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
          ########## ######---- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 1     TGTTCTCGCG CAG------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 2     TGTTCTCGCG CAG------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 4     CACTCATTCG GAG------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 5     CACTCATTCG GAG------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 7     GAACCACAAG GAC------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 9     TGTCCACGCC GCG------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 11    TAACGATTAT AGA------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 13    AAATGTTGCG CAT------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 15    CAACAACAAG GAC------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 17    TTATGAGGAT CCTCAGGTAA GTTTTGGTGT TGTTTGAGGG ATGGGGCAAG GGGTTGTCTG TCGTGAACAA CAAAAGGGGC ACGGAACAAA TGCTAACGCC
SEQ 18    TTATGAGGAT CCTCAG---- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 20    AAACTCGGTG GCC------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 21    AAACTCGGTG GCC------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 23    TAATCACAAG GAT------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 25    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 26    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 28    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 29    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 32    TATCCATCCT AAG------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 34    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 36    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 37    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 39    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 41    GAATCATCCT CAG------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 43    CTATGAGAGT TTT------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 82    CATCCACAAG ATG------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 84    CCTGGTTCCA TTC------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------

          1701       1711       1721       1731       1741       1751       1761       1771       1781       1791
          ---------- -10------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
          ------#### ########## ########## ########## ##-------- ---------- ---------- ---------- ---------- ----------
SEQ 1     ---------- ---------- -------CAG AAGATCAAGT CCGGCCCTGC CTTCCAGGTG CCTTTTGCCG TGGCCGTGAA GAAGGCCGTC GGCGAC-----
SEQ 2     ---------- ---------- -------CAG AAGATCAAGT CCGGCCCTGC CTTCCAGGTG CCTTTTGCCG TGGCCGTGAA GAAGGCCGTC GGCGAC----
SEQ 4     ---------- ---------- -------CAG CATATCCACG CGAAGCCAGG CTTCCAGGCA CCCTTTGCTA TTGCCGTCAA GAACGCCGTC GGGGAC----
SEQ 5     ---------- ---------- -------CAG CATATCCACG CGAAGCCAGG CTTCCAGGCA CCCTTTGCTA TTGCCGTCAA GAACGCCGTC GGGGAC----
SEQ 7     ---------- ---------- -------CAG AAGATCAACC TGCACACGGC CTACCAGACG GACCTGGCCG GGCAGATTCG CCAGGCCATC CGAGCG----
SEQ 9     ---------- ---------- -------CAG AAGATCAAGT CCGGGCCGGC TTTCCAGGCT CCCTTCGCTG TGGCTATCAA GAAGGCCGTT GGCGAT----
SEQ 11    ----------C AACCACCAAG ATCTGGGATC AGTAAAGAGT TGAGAGAGCC AATCCATGTT CCGTTGTCTC GTGCAATTAA ACAACATGTT GGTGAC----
SEQ 13    ----------T GCAAATCTAG ATATCTATTA AATGACGACA AACAACTACC TTCTCAAGTG CCCTTGGCTC GTAAATTGAA AAGCCACATT AGAAAC----
SEQ 15    ---------- ---------- -------CAG AAGATCAACG TCCACACCTA CTACCAGATC GACATGGCCG AGCAGATCCG CGCGGCCGTG CACGAGGCCG
SEQ 17    ATACAGATGG CCAACGGTCC CAAGCCCGAA AAGTCCGAAC GCACCATGGC CCGCGAGGCC TTCTTCCTCG AGTTCGCCAA GATCATCCGC ACCAAG---T
SEQ 18    ------ATGG CCAACGGTCC CAAGCCCGAA AAGTCCGAAC GCACCATGGC CCGCGAGGCC TTCTTCCTCG AGTTCGCCAA GATCATCCGC ACCAAG---T
SEQ 20    ---------- ---------- -------CAA AAGATCGAGC TCACGCCGTA CTACCAGATC GACCTGGCAG CCAAGATCCG CGAGGCCGTC GGCGAT----
SEQ 21    ---------- ---------- -------CAA AAGATCGAGC TCACGCCGTA CTACCAGATC GACCTGGCAG CCAAGATCCG CGAGGCCGTC GGCGAT----
SEQ 23    ---------- ---------- -------CAA AGAATTGAGG TGAAGGATTG CTATCAAGTT CCTTTTGCGG AAAAGATTAA GGATCAAGTG AATGGA----
SEQ 25    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 26    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 28    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 29    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 32    ---------- ---------- -TCCGCCATC GCCATCAAGT CCGGTCCTGC TTACCAGGTA GACCTCGCCA AACAGGTAAA GAAGGCTGTT GGCGAT----
SEQ 34    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 36    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 37    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 39    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 41    ---------- ---------- -------CAG AAAATCAACA TGTTCAACAC C_---------- ---------- ---------- ---------- ----------
SEQ 43    ---------G GTTTTGCGCA CCGCAAGGAG TCCAGCCGCA AGCGGGAGAA CTATTTTATC GAGTTCGCCG AGGTCATCCG CAAGGCCGTC AAGCAC-----
SEQ 82    ---------- ---------- -------CAA AAGGTCGCTG CTGGTCCCGG TTACCAGGCA CCTCTTGCCA AGGCGATCAA GAAGTCAGTT GGAGAC-----
SEQ 84    ---------- ---------- -------CAA AAAATCACCG TGGGAGCCGG ATACCAGCTA TTCGGAGCAA AAGCCGTTCG CGATGCTCTG GCCAAA-----
```

```
          1801       1811       1821       1831       1841       1851       1861       1871       1881       1891
          ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
          ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 1     -----AAGCT GCTGGTTGCC GCCGTGGGTG CCATCACC-- ---------- ---------- ---------- ---------- ------AACG GCAAGCAGGC
SEQ 2     -----AAGCT GCTGGTTGCC GCCGTGGGTG CCATCACC-- ---------- ---------- ---------- ---------- ------AACG GCAAGCAGGC
SEQ 4     -----AAACT CGCAGTGGCA TCAGTGGGTA TGATTGCC-- ---------- ---------- ---------- ---------- ------AGCG CGCATTTGGC
SEQ 5     -----AAACT CGCAGTGGCA TCAGTGGGTA TGATTGCC-- ---------- ---------- ---------- ---------- ------AGCG CGCATTTGGC
SEQ 7     -----GCTGG CGCGTCGACT CTTGTGGGTG CTGTAGGTCT GATCACCGAT TCGGAACAGG CGAGGGGACT AGTTCAGGGA GCGGACGAGG CGACTGCAGC
SEQ 9     -----AAGCT CCTTGTTGCG ACGGTGGGCA CGATCACG-- ---------- ---------- ---------- ---------- ------AACG GTAAGCAGGC
SEQ 11    -----AAGTT ATTGGTCAGT TGCGTTGGTG GGCTTGAA-- ---------- ---------- ---------- ---------- ---------A AAGATCCTGA
SEQ 13    -----CGATG TTTGATCGCA TGCAGTGGAG GATTAGAT-- ---------- ---------- ---------- ---------- ---------C GAGACATATT
SEQ 15    GCAAGCAGCT CCTCGTCGGT GCCGTCGGCT TGGTCACC-- -------TCG GCTGAGATCG CCAAGGAGAC CGTCCAGGAG AAGGAGGATG GCAGAGTCAC
SEQ 17    TCCCCAAGCT TCCTCTCATG GTCACCGGCG GCTTCCGC-- ---------- ---------- ---------- ---------- ------ACTC GTCAGGGCAT
SEQ 19    TCCCCAAGCT TCCTCTCATG GTCACCGGCG GCTTCCGC-- ---------- ---------- ---------- ---------- ------ACTC GTCAGGGCAT
SEQ 20    -----AGGTT GCTCATAGGC GCGGTCGGCA ACATCAAC-- ---------- ---------- ---------- ---------- ------ACGG CTGACATTGC
SEQ 21    -----AGGTT GCTCATAGGC GCGGTCGGCA ACATCAAC-- ---------- ---------- ---------- ---------- ------ACGG CTGACATTGC
SEQ 23    --------AT ACTACTTGGC GCTGTCGGAA TGATCAGG-- ---------- ---------- ---------- ---------- ------GATG GTCTTACGGC
SEQ 25    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 26    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 28    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 29    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 32    -----AGTGT ACTTGTTTCA GCAGTAGGTG GAATCAAG-- ---------- ---------- ---------- ---------- ---------A CTGGACATCT
SEQ 34    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 36    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 37    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 39    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 41    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 43    -----ATGGT GGTCTACACC ACCGGCGGCT TCAAGACG-- ---------- ---------- ---------- ---------- ---GTGGGCG CCATGGTCGA
SEQ 82    -----AAGAT GTTGATCAGC ACTGTTGGTA GCATCAAG-- ---------- ---------- ---------- ---------- ------ATAG GTACCCTTGC
SEQ 84    --ATCGAACC CGACGCGTCC AAACGCATGC TCGTCGGGG- ---------- ---------- ---------- ---------- ----CCGTGG GAATGATGGA

          1901       1911       1921       1931       1941       1951       1961       1971       1981       1991
          ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
          ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 1     GAATCAG--- ---ATTCTAG AGGAGCAG-- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 2     GAATCAG--- ---ATTCTAG AGGAGCAG-- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 4     CAATTCC--- ---TTGTTGG AGAAGGAC-- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 5     CAATTCC--- ---TTGTTGG AGAAGGAC-- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 7     CGAGGCAATG CTGTCGGGAC CTGAACCC-- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 9     GAACAAG--- ---CTGCTTG AGGAGGAG-- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 11    ATTGCTCAAC AAATATTTAG AAGAAGGA-- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 13    TAAACTCGAT GAGTTTATTG CTAATGGT-- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 15    CATCCAGCGC GAGAACGGCG CCAAGACT-- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 17    GGAGGCC--- ---GCTTTGG AATCCGAT-- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 18    GGAGGCC--- ---GCTTTGG AATCCGAT-- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 20    GCGCGATGTC GTGGATGAGC AGGGCGCCGA GAAGGTGGCC GAGGCCAAGC AGACGCATGA CACCATCGAG GTCGTGAGCG AATCACATGG CGGCAAGACC
SEQ 21    GCGCGATGTC GTGGATGAGC AGGGCGCCGA GAAGGTGGCC GAGGCCAAGC AGACGCATGA CACCATCGAG GTCGTGAGCG AATCACATGG CGGCAAGACC
SEQ 23    GAATGAAATC CTAGAAAGTG GAAAAGCT-- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 25    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 26    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 28    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 29    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 32    TGCTGAA--- ---GAGGTTT TGCAATCT-- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 34    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 36    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 37    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 39    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 41    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 43    CGCGCTGCAG GGCGTCGATG GG-------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 82    GGAGGAG--- ---ATCATCG CTGGAGGAGA GGACGATACC ---------- ---------- ---------- ---------- ---------- ----------
SEQ 84    AGGTTCC--- ---TACGATT CGCCCAAC-- ---------- ---------- ---------- ---------- ---------- ---------- ----------

          2001       2011       2021       2031       2041       2051       2061       2071       2081       2091
          ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
          ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 1     GATATCGACG TTGCGCTGGT TGGCCGTGGG TTCCAGAAGG ATCCCGGTCT GGCCTGGACG TTTGCTCAGC ACCTCGGCGT C--------- ----------
SEQ 2     GATATCGACG TTGCGCTGGT TGGCCGTGGG TTCCAGAAGG ATCCCGGTCT GGCCTGGACG TTTGCTCAGC ACCTCGGCGT C--------- ----------
SEQ 4     GGACTGGACC TTGTGCTGGT TGGACGTGGC TTCCAGAAGA ACCCGGGGCT GGTGTGGGCG TGGGCCGACG AGCTGAATGT A--------- ----------
SEQ 5     GGACTGGACC TTGTGCTGGT TGGACGTGGC TTCCAGAAGA ACCCGGGGCT GGTGTGGGCG TGGGCCGACG AGCTGAATGT A--------- ----------
SEQ 7     AAGGCGGATG CCATTCTGAT AGCCCGTCAG TTCCTGCGCG AGCCAGAATG GGTGTTTTCC ACGGCGAGAA AGTTGGGCGT G--------- ----------
SEQ 9     GGATTGGATG TTGCGCTTGT GGGACGTGGT TTCCAGAAGG ATCCCGGTCT GGCGTGGACT TTCGCGCAGC ATCTTGATGT T--------- ----------
SEQ 11    ACATTTGATC TTGCTTTGAT CGGTAGAGGA TTTTTAAGAA ATCCAGGTTT GGTATGGGAG TTTGCCGATA AACTTGGTGT T--------- ----------
SEQ 13    GACTTTGATA TAGCATTGAT AGGTAAAGGA TTTCTCAAAA ACACTGGATT GATCAGCCGT ATTGCTGACC AATTGCAAGC A--------- ----------
SEQ 15    CGTGCCGATA TGGTCCTTGT TGCCAGGCAG TTCTTGAAGG AGCCCGAGTT CGTCCTCACT GTCGCCGACG AGTTGGGTGT T--------- ----------
SEQ 17    GATTGCGACA TGATCGGTAT CGGACGCCCG GCCATCATCA ACCCTTCGCT TCCCGCCAAC TTGATCCTCA ACCCGGAGGT G--------- ----------
SEQ 18    GATTGCGACA TGATCGGTAT CGGACGCCCG GCCATCATCA ACCCTTCGCT TCCCGCCAAC TTGATCCTCA ACCCGGAGGT G--------- ----------
SEQ 20    AAGGCGGATC TGGTCCTCAT TGCTCGCCAG TTCCTGCGCG AGCCTGAGTT TGTGCTGAGG ACGGCGCATA ACCTTGGGGT C--------- ----------
SEQ 21    AAGGCGGATC TGGTCCTCAT TGCTCGCCAG TTCCTGCGCG AGCCTGAGTT TGTGCTGAGG ACGGCGCATA ACCTTGGGGT C--------- ----------
SEQ 23    ------GATG TTACTTTTGT CGCAAGGGAG TTCTTAAGGA ACCCGTCGTT GGTGCTAGAC AGCGCGAACC AGTTGGGTGA A--------- ----------
SEQ 25    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 26    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 28    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 29    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 32    GGTATCGACA TTGTGAGGGC TGGACGTTGG TTCCAACAGA ATCCTGGTCT GGTTCGAGCT TTTGCTAACG AGCTTGGCGT G--------- ----------
SEQ 34    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 36    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 37    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 39    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 41    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 43    ---------- --ATAGGCAT CGGGCGCGCA GCCGGTTCGG AGCCGGACCT CGCCAAGGAC ATCATCGCGG GCAAGGTGTC CAGCATTATC AAATACGCCA
SEQ 82    CCCTTGGATC TTGTGGCTTC AGGCCGTCTG TTCCAGAAGA ACACTGGACT TGTTTGGTCA TGGGCTGACG ATCTGAACAC T--------- ----------
SEQ 84    GGCCAAGACC GCAGCCAGAT TGGCAAGTTG GCCGAGCAGT CGATTCAGAG CGGAGAGTGT GATGCGGTAC TGTTGGCACG T--------- ---GGATTGA
```

```
          2101       2111       2121       2131       2141       2151       2161       2171       2181       2191
          ---------- ---------- ---------- ---------- --11------ ---------- ---------- ---------- ---------- ----------
          ---------- ---------- ---------- --******** ********** ***------- ---------- ---------- ---------- ----------
SEQ 1     ---------- ---------- ------GAAA TCTCCATGGC CAACCAGATC CGCTGGGGCT TCACCCGGCG TGGAGGCACC CCGTACATTG ATCCTTCGGT
SEQ 2     ---------- ---------- ------GAAA TCTCCATGGC CAACCAGATC CGCTGGGGCT TCACCCGGCG TGGAGGCACC CCGTACATTG ATCCTTCGGT
SEQ 4     ---------- ---------- ------GAGA TCTCCATGGC TAATCAGATC CGATGGGGTT TCTCGCGGCG CGGTGCTGGT CCTTACCTCA GGAAGAAACT
SEQ 5     ---------- ---------- ------GAGA TCTCCATGGC TAATCAGATC CGATGGGGTT TCTCGCGGCG CGGTGCTGGT CCTTACCTCA GGAAGAAACT
SEQ 7     ---------- ---------- ------CCGG TGACTGTCCC GGTGCAGTTT GGCAGGGCCA TTTAG----- ---------- ---------- ----------
SEQ 9     ---------- ---------- ------GAGA TTGCGATGGC GAGTCAGATT CGGTGGGGAT TCACAAGGCG CGGGGGCACG CCTTATATCG ACCCCAAAGC
SEQ 11    ---------- ---------- ------AGAC TCCACCAGGC CTTGCAGTTA GGTTGGGGTT TCTGGCCCAA CAAACAACAA ATTGTTGATT TGATTGAAAG
SEQ 13    ---------- ---------- ------CAAT TCAGAACAGC ACCTCAATAT AAGTTGGCCT TATCATAA-- ---------- ---------- ----------
SEQ 15    ---------- ---------- ------GATG TCAAGGCCCC TGTTCAGTAC CTCCGTGGTC CTCTTAGCAG CAGGCCCAAG AAGTTGACCA CTGTTCCTTA
SEQ 17    ---------- ---------- ------CCGG ATGCGGATGC CCGCTTGTTC GACAAGAAGA GGGCTGAGCC GCACTGGATC GTTGAGAAGT TGGGCATGAA
SEQ 18    ---------- ---------- ------CCGG ATGCGGATGC CCGCTTGTTC GACAAGAAGA GGGCTGAGCC GCACTGGATC GTTGAGAAGT TGGGCATGAA
SEQ 20    ---------- ---------- ------AATG TGCAGTGGCC TCACCAATAC CACAGAGCAG TGTGGCGCAA GGGTGCAAGG ATTTGA----- ----------
SEQ 21    ---------- ---------- ------AATG TGCAGTGGCC TCACCAATAC CACAGAGCAG TGTGGCGCAA GGGTGCAAGG ATTTGA----- ----------
SEQ 23    ---------- ---------- ------AATG TTGCATGGCC AGTTCAGTAT GACTATGCAG TTAAGGGACA CAGAAAGTTA CGTTGA----- ----------
SEQ 25    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 26    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 28    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 29    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 32    ---------- ---------- ------GAGG TCAAGATGGC GAACCAGATT GATTGGAGCT TCAAGGGACG TGGAAAGAAA GTGAACAAGA GTTCTTTATA
SEQ 34    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 36    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 37    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 39    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 41    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 43    TGGGGGAGGA CGAGTTTGTG CTGCAGTTGA CTGCCTGCTC GGCGCAAATA AGGCTGATGG CCAAGGGCGA GGAGCCGTTT GAC------- ----------
SEQ 82    ---------- ---------- ------TCTA TCCAGATCGC TCATCAGATC GCATGGGGTT TCGGTGGCAG AGCTAAGAAG AACGCTCCCA AGCTTGTCTT
SEQ 84    TGTCCTACCC AAGCTGGACC GAGGATGCTA GTGTAGCGCT GATGGGTACC AGGGCAGCTG GCAACCCGCA GTACCATCGC GTTCACGTGG CTAAGAAGTG

          2201       2211       2221       2231       2241       2251       2261       2271       2281       2291
          ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
          ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 1     GTACAAGCAG TCTATTTTCG ATGTATAG-- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 2     GTACAAGCAG TCTATTTTCG ATGTATAG-- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 4     CGAGAAGATA TAA------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 5     CGAGAAGATA TAA------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 7     ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 9     TTATAAGGAG AGCATCTTTG AGTAA----- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 11    AACATCTAAA TTAGAAGTAA ATTAG----- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 13    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 15    A--------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 17    GTCCATTGTT GGTGCTGGTG TTGAGGTGGT ACGTCACGTT CCAACCCCAT TTGCTTCATT GTGTTTCCGA GTATGTCATG CTGACTTGGT TCTTTTCTAG
SEQ 18    GTCCATTGTT GGTGCTGGTG TTGAGGTG-- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 20    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 21    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 23    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 25    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 26    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 28    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 29    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 32    G--------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 34    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 36    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 37    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 39    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 41    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 43    ---------- ---------- ---------- ---------- ---------- -----ATCTC AAACGCCGAC GAGGTGGCGC GGGTGACGCA GTTGATGGCG
SEQ 82    A--------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 84    A--------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------

          2301       2311       2321       2331       2341       2351       2361       2371       2381       2391
          ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
          ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 1     AGTATAGATA GAGTTGAAGA TGATACCTCA TAGACGATCA ATGGACCCTT GCATATTATT TCTCGTCTCC TGCGTATGTT CAAGGTATTC ACAGTAGCTG
SEQ 2     AGTATAGATA GAGTTGAAGA TGATACCTCA TAGACGATCA ATGGACCCTT GCATATTATT T--------- ---------- ---------- ----------
SEQ 4     ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 5     ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 7     ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 9     ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 11    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 13    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 15    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 17    ACGTGGTATG TGAGCGAGCT CAAGAAGCTG GCCAAGTTTT AG-------- ---------- ---------- ---------- ---------- ----------
SEQ 18    ACGTGGTATG TGAGCGAGCT CAAGAAGCTG GCCAAGTTTT AG-------- ---------- ---------- ---------- ---------- ----------
SEQ 20    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 21    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 23    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 25    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 26    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 28    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 29    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 32    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 34    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 36    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 37    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 39    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 41    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 43    GAGGGCAAGG TG-------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 82    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
SEQ 84    ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
```

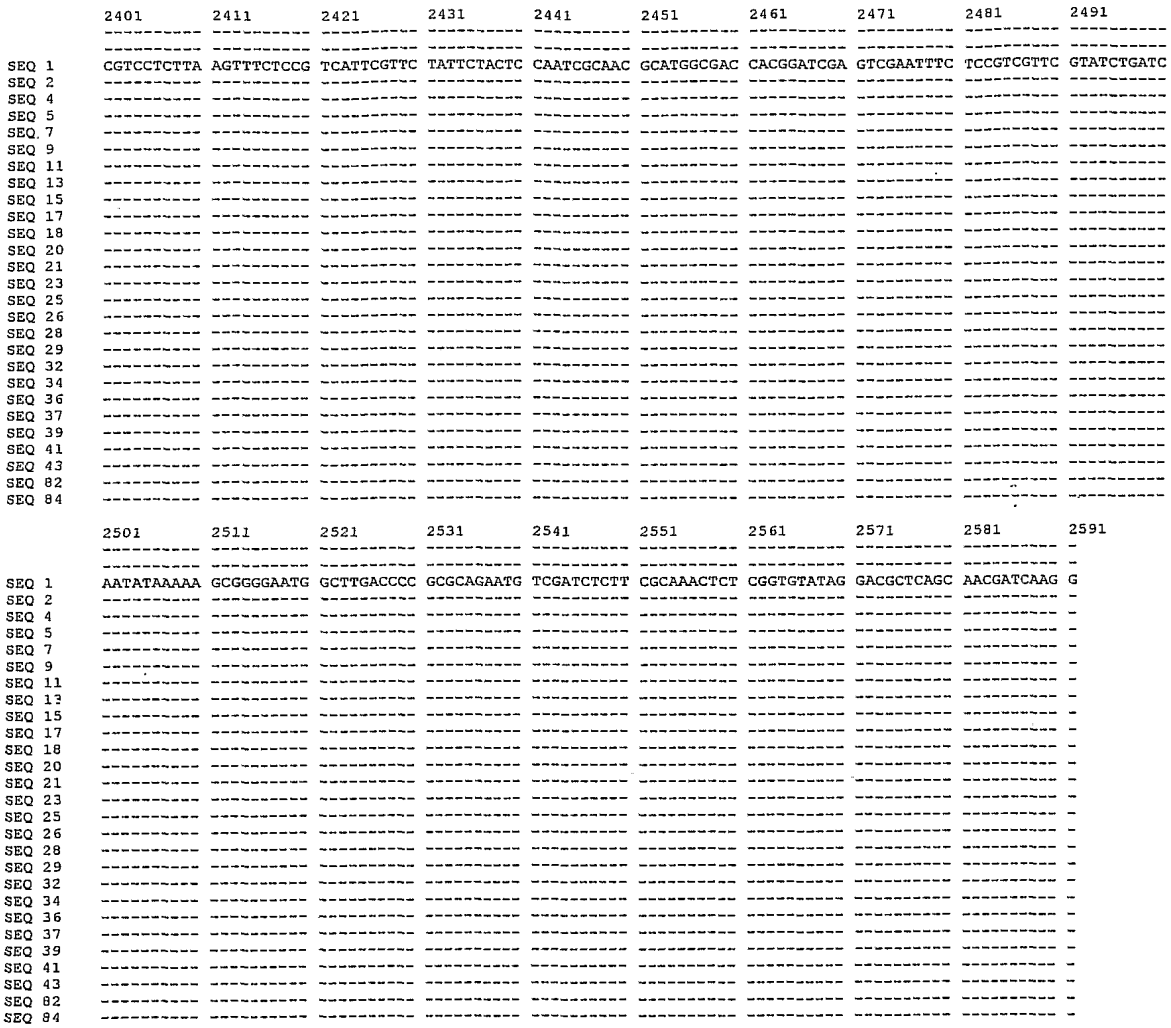

Figure 2. A multiple alignments of the 2031 OR nucleic acid sequence from *A. fumigatus* (SEQ 1,2) along with related 2031 ORs from other fungi and bacteria (see also Example 4). Regions 1-11, marked with * or #, refer to regions conserved at the amino acid level between Ors but not OYEs.

Fungal 2031 ORs are given by SEQ ID No.: SEQ ID Nos. 1, 2, 4, 5, and 7, *A. fumigatus;* SEQ ID No. 9, *A.nidulans;* SEQ ID Nos. 11 and 13, *C. albicans;* SEQ ID Nos. 15, 17 and 18, *N. crassa;* SEQ ID Nos. 20, 21 and 43, *M. grisea;* SEQ ID No. 23 (NP_595868), *S. pombe;* SEQ ID Nos. 25 and 26, *C. trifolii;* SEQ ID Nos. 28, 29, 31, 32 and 34, *F. sporotrichioides;* SEQ ID Nos. 36, 37 and 82, *F. graminearum;* SEQ ID Nos. 39 and 41, *M. graminicola;* SEQ ID No. 84, *U. maydis.*

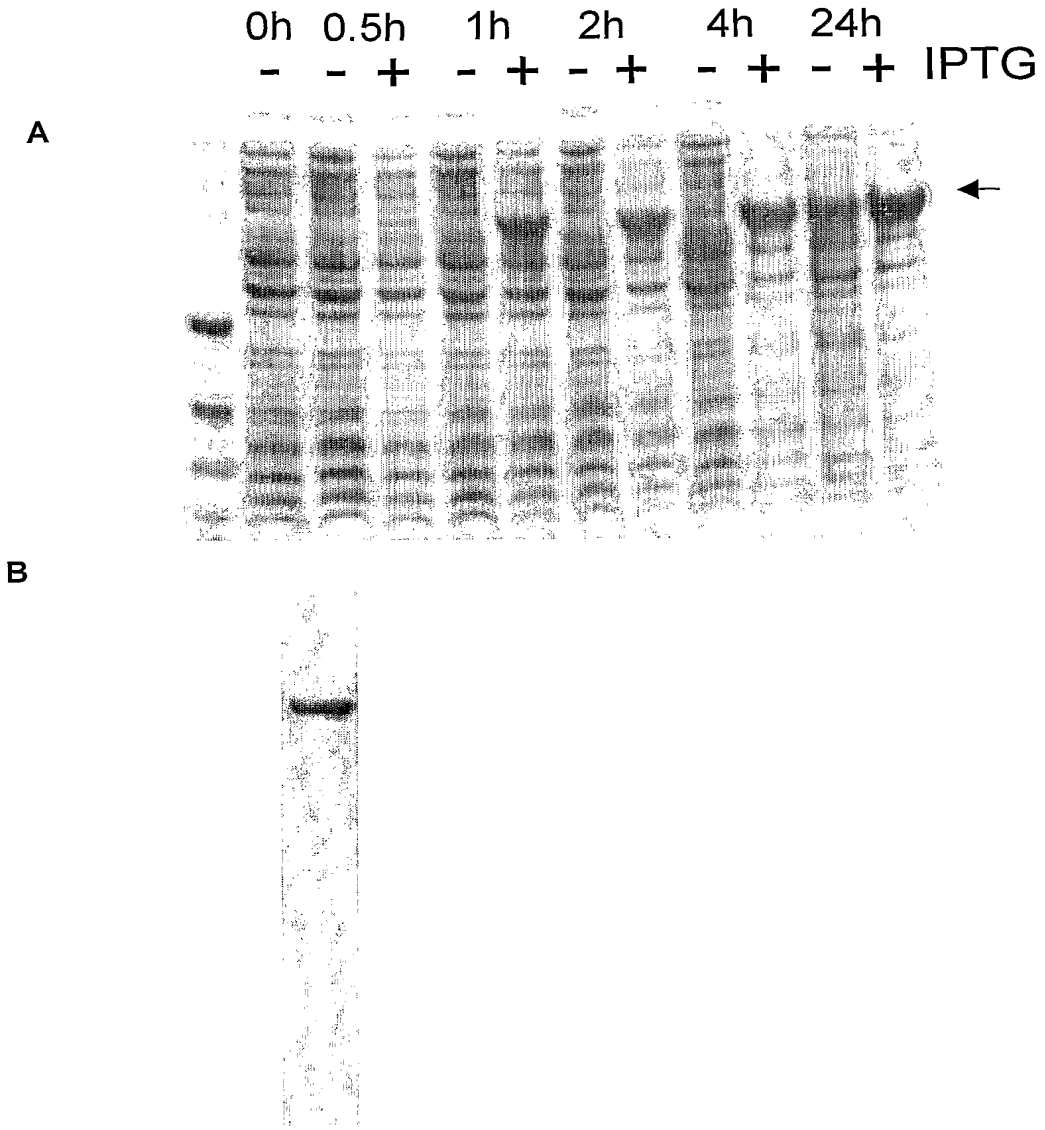

Figure 3. Recombinant 2031 OR. (A) Time course of recombinant 2031 OR induction over 24 hours after the addition of IPTG (samples without IPTG are also shown). The gel was stained with coomassie; A prominent band of the correct molecular weight (marked with an arrow) is seen. (B) Coomassie stained gel showing purified recombinant 2031.

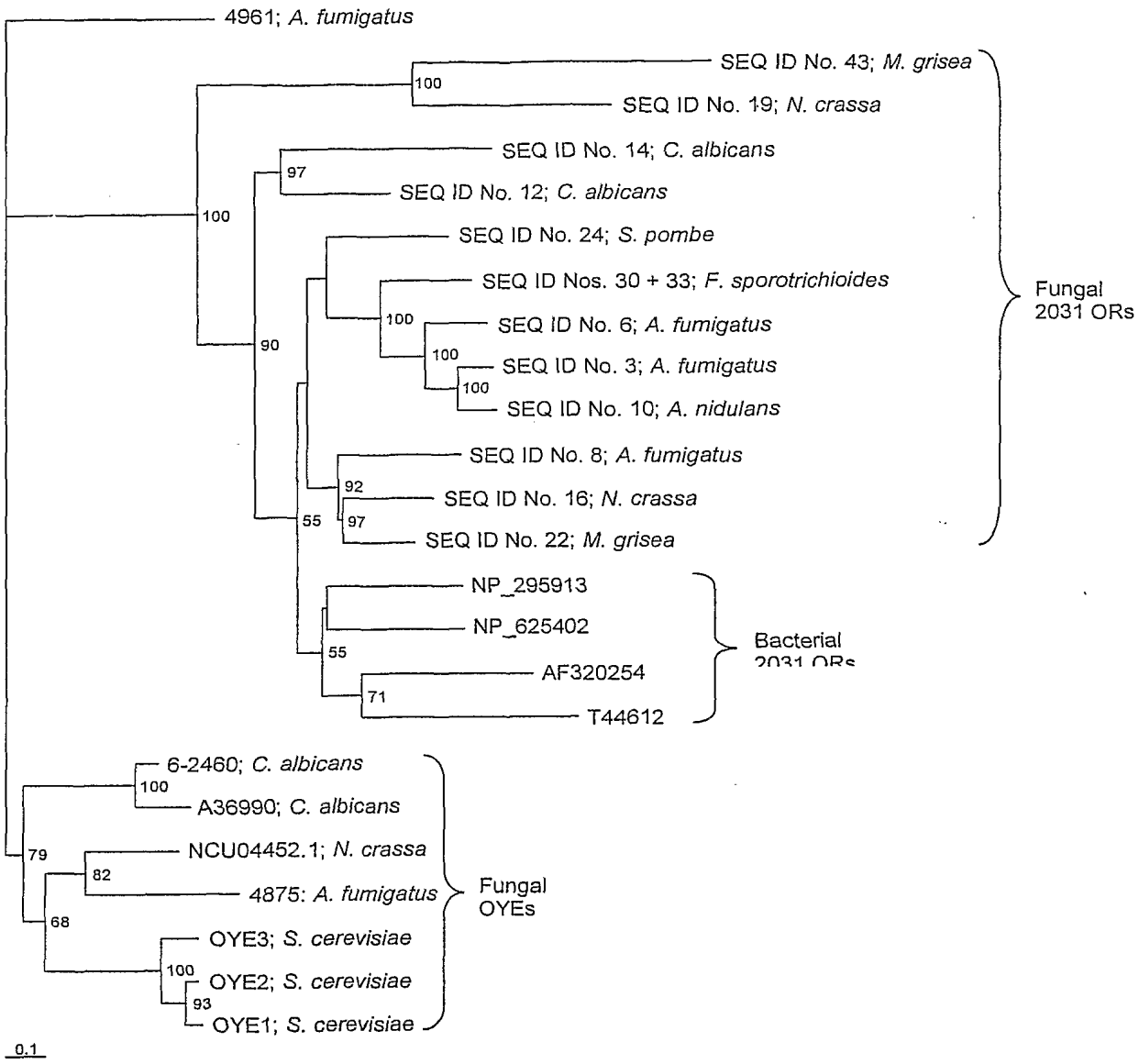

Figure 4. Phylogenetic tree showing relationships between *A. fumigatus* 2031 OR and similar proteins. This demonstrates a 2031 OR clade, which can be distinguished from the OYE proteins.

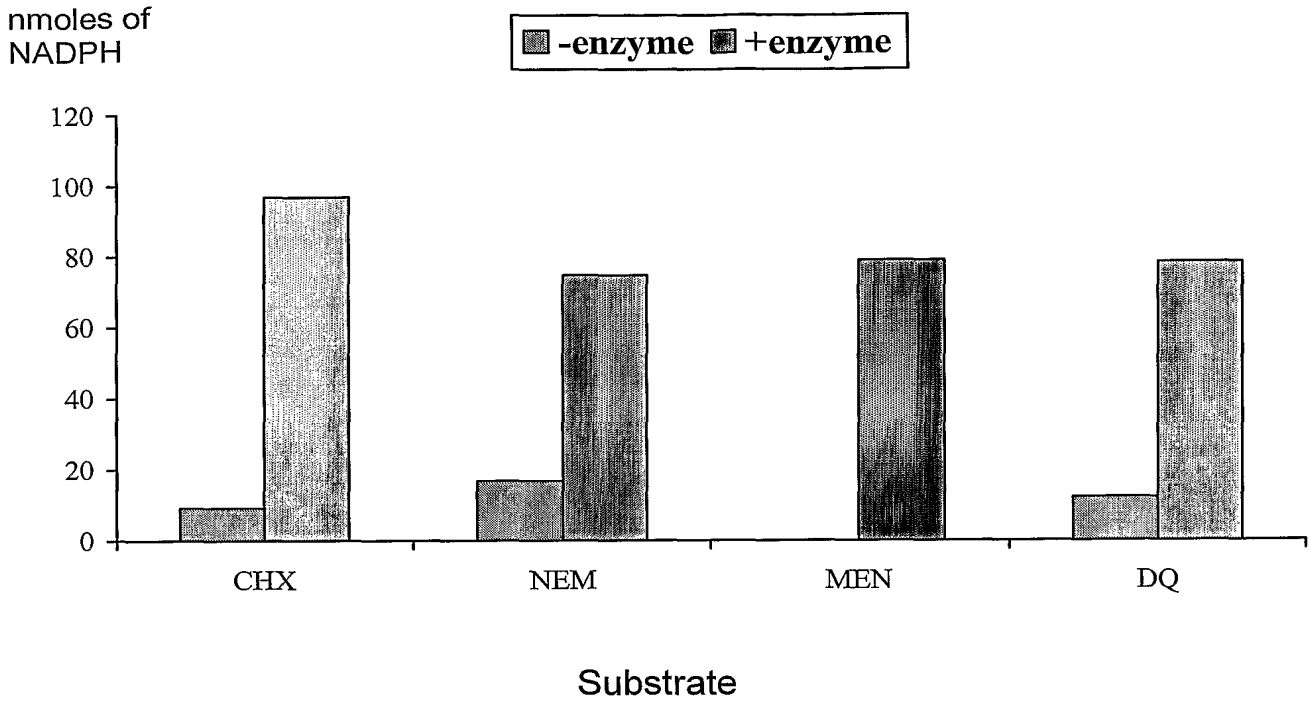

Figure 5: NADPH dehydrogenase activity of recombinant 2031 OR with cyclohexenone (CHX), N-ethylmaleimide (NEM), menadione (MEN) or duroquinone (DQ) as substrates. Final concentrations in the assay were as follows: 500 μM substrate, 120 μM NADPH, 1 μg/200 μL 2031 OR.

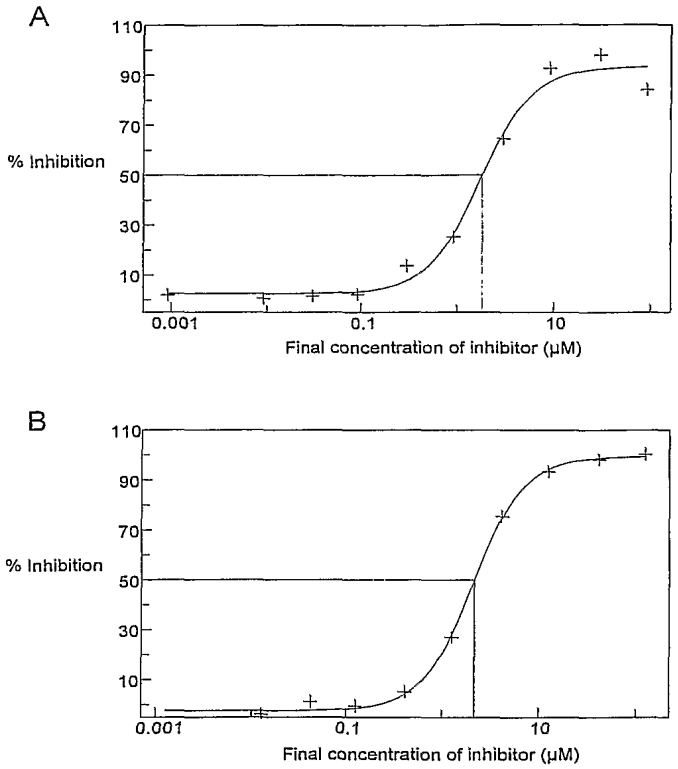

Figure 6: Inhibition of 2031 OR function by two inhibitors (shown in A and B) identified by high-throughput screening.